

salesforce

Containing an Elephant

Hadoop/HBase on Kubernetes

Dhiraj Hegde, Architect at Salesforce

@dhirajtxt | dhegde@salesforce.com



Why Kubernetes?



Why Kubernetes?



- Immutability of Containers (Pods)

Why Kubernetes?



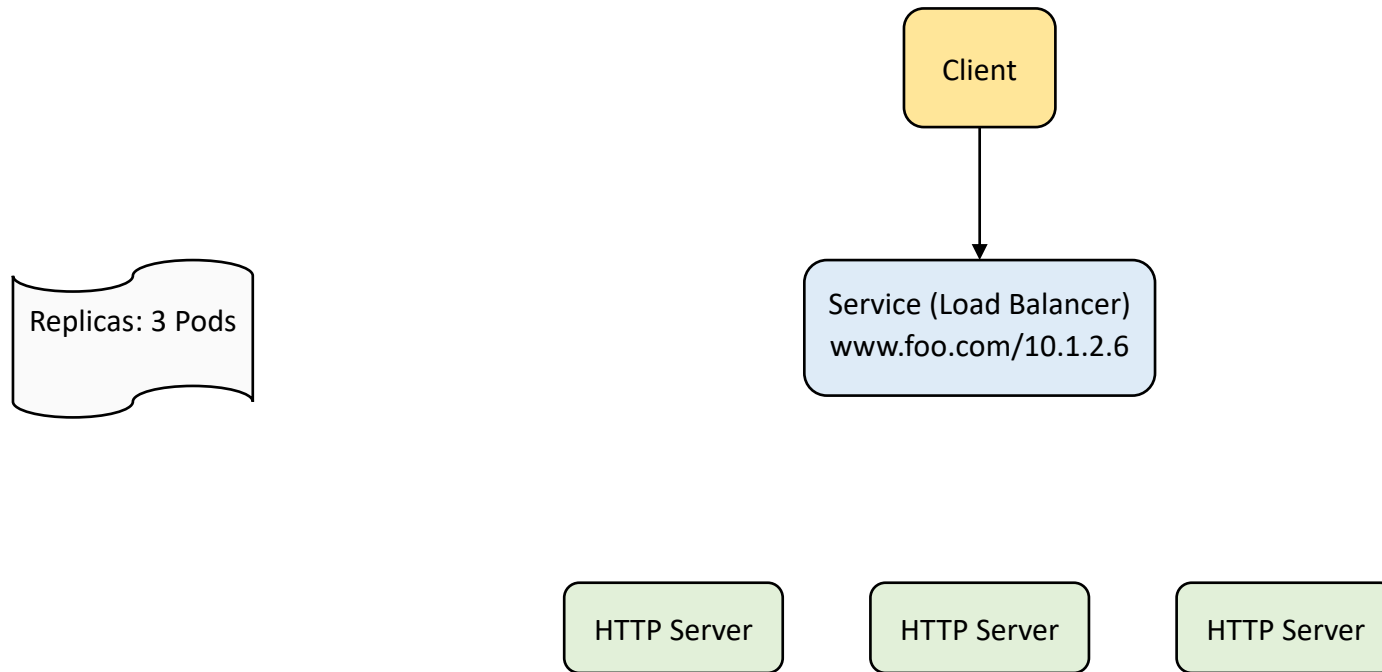
- Immutability of Containers (Pods)
- Availability and Scalability

Why Kubernetes?



- Immutability of Containers (Pods)
- Availability and Scalability
- Consistency across clouds

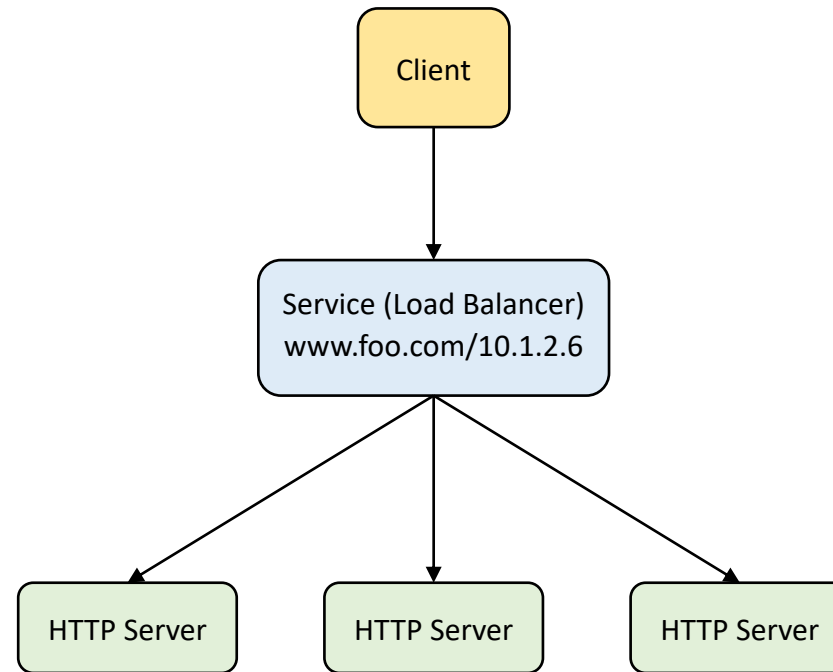
Kubernetes and Stateless Applications



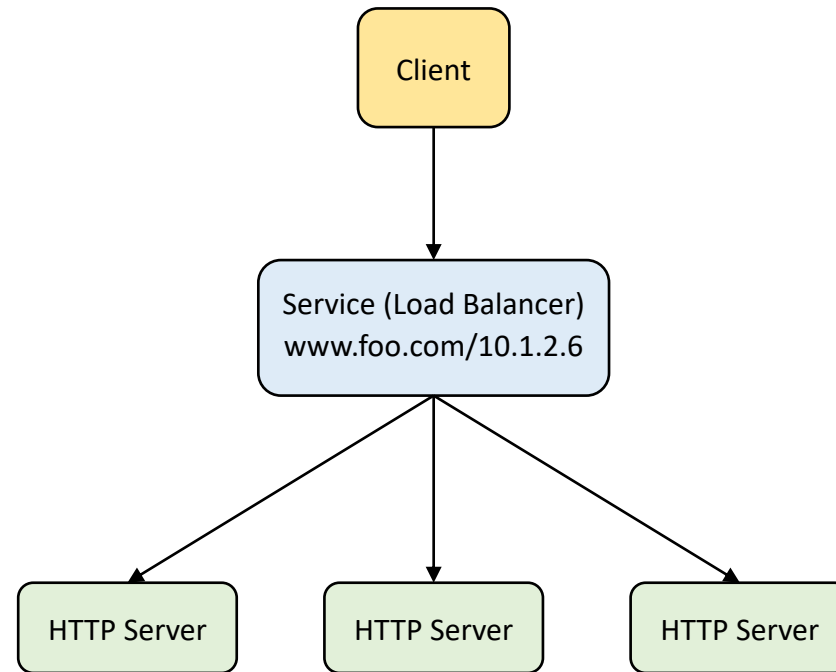
Kubernetes and Stateless Applications



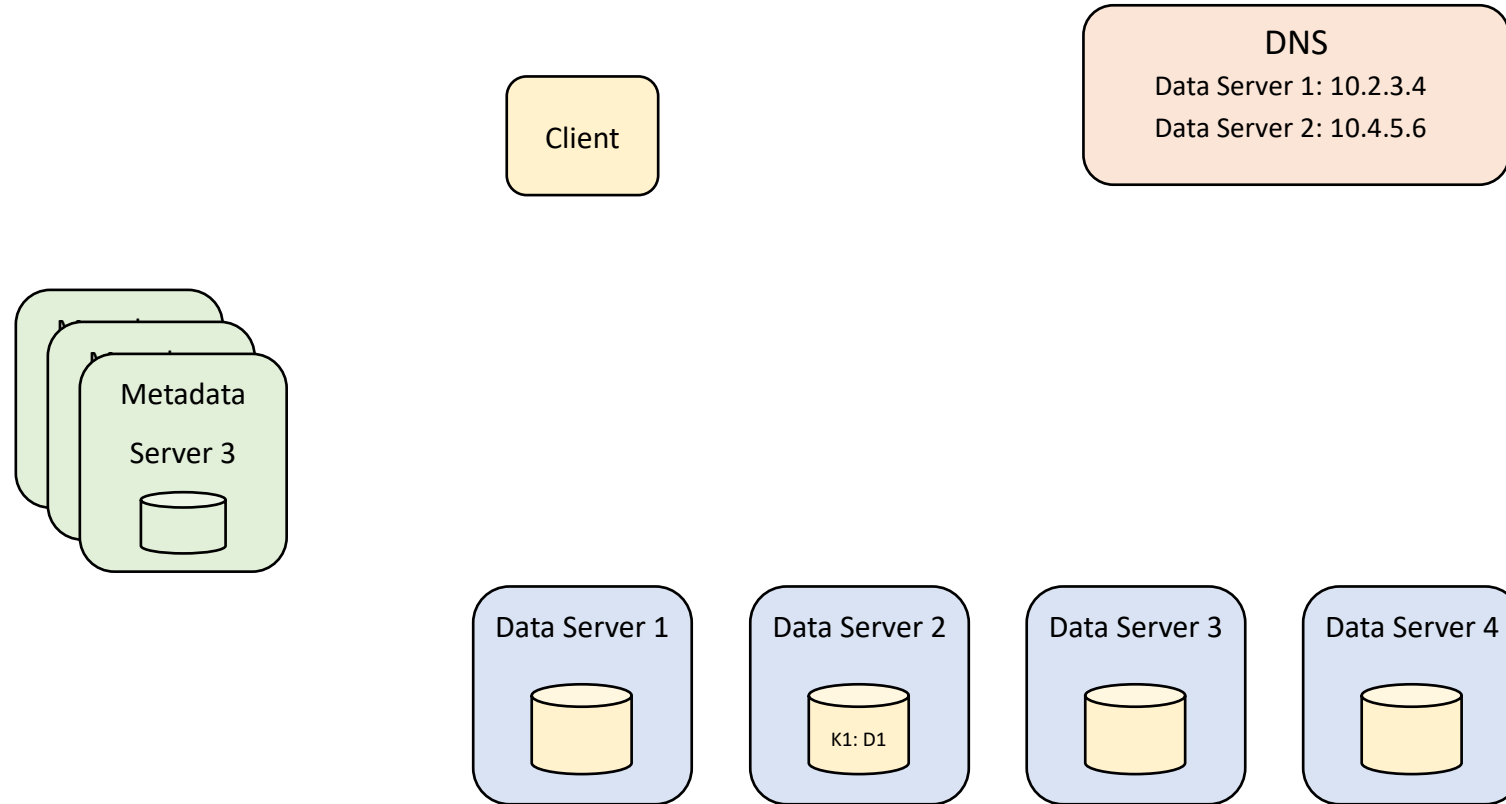
Replicas: 3 Pods



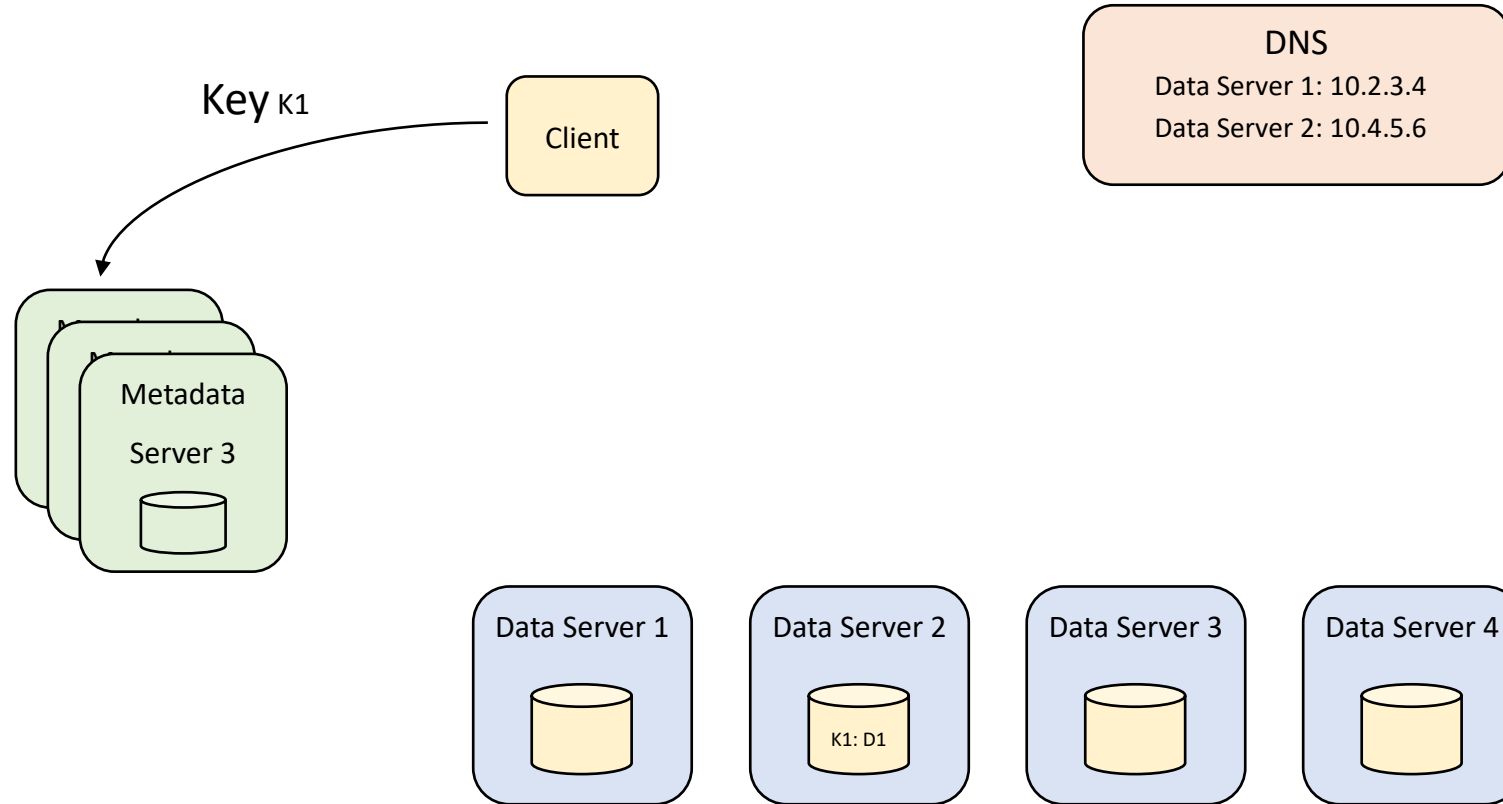
Kubernetes and Stateless Applications



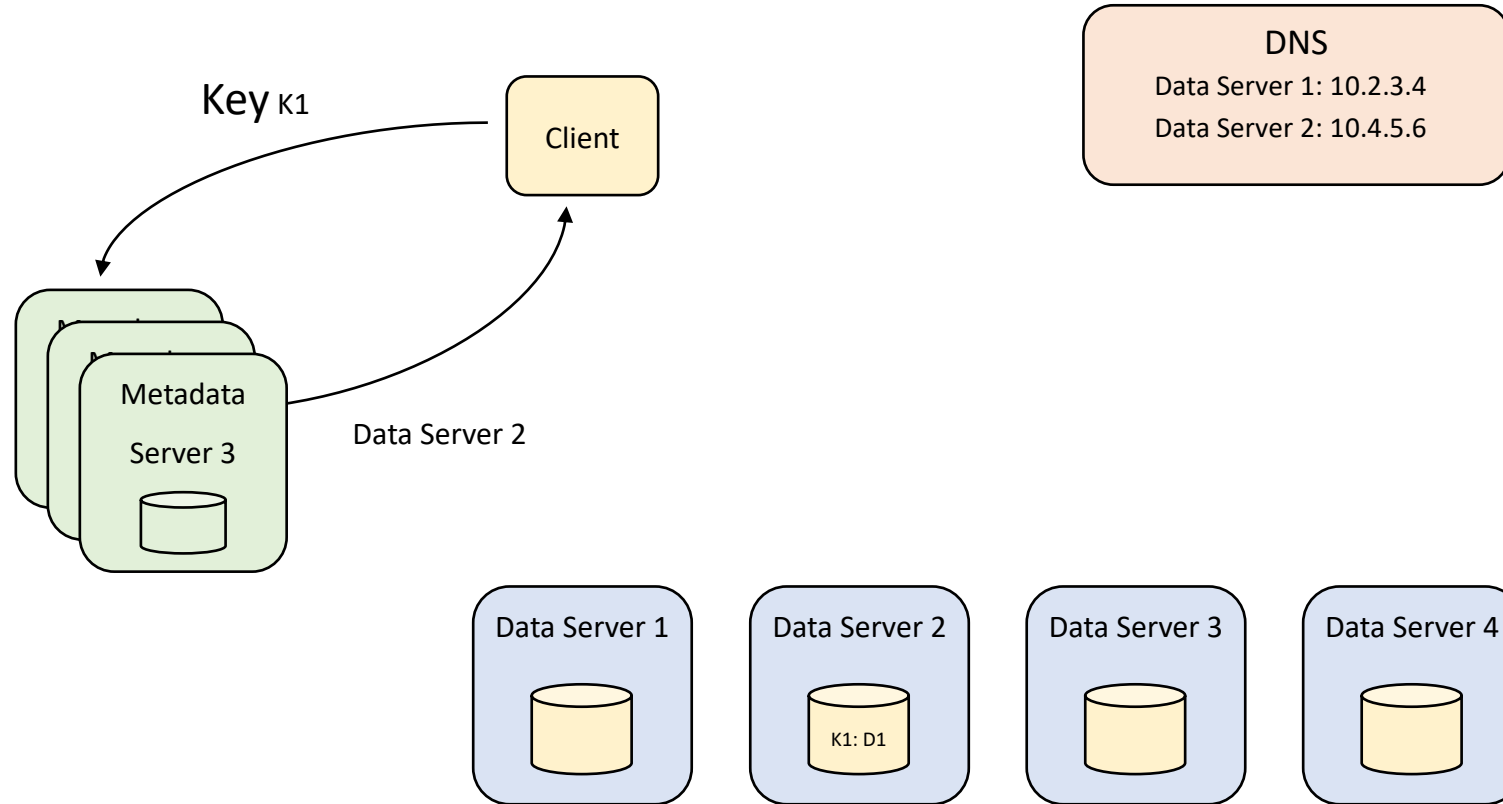
Hadoop/HBase Stateful Applications



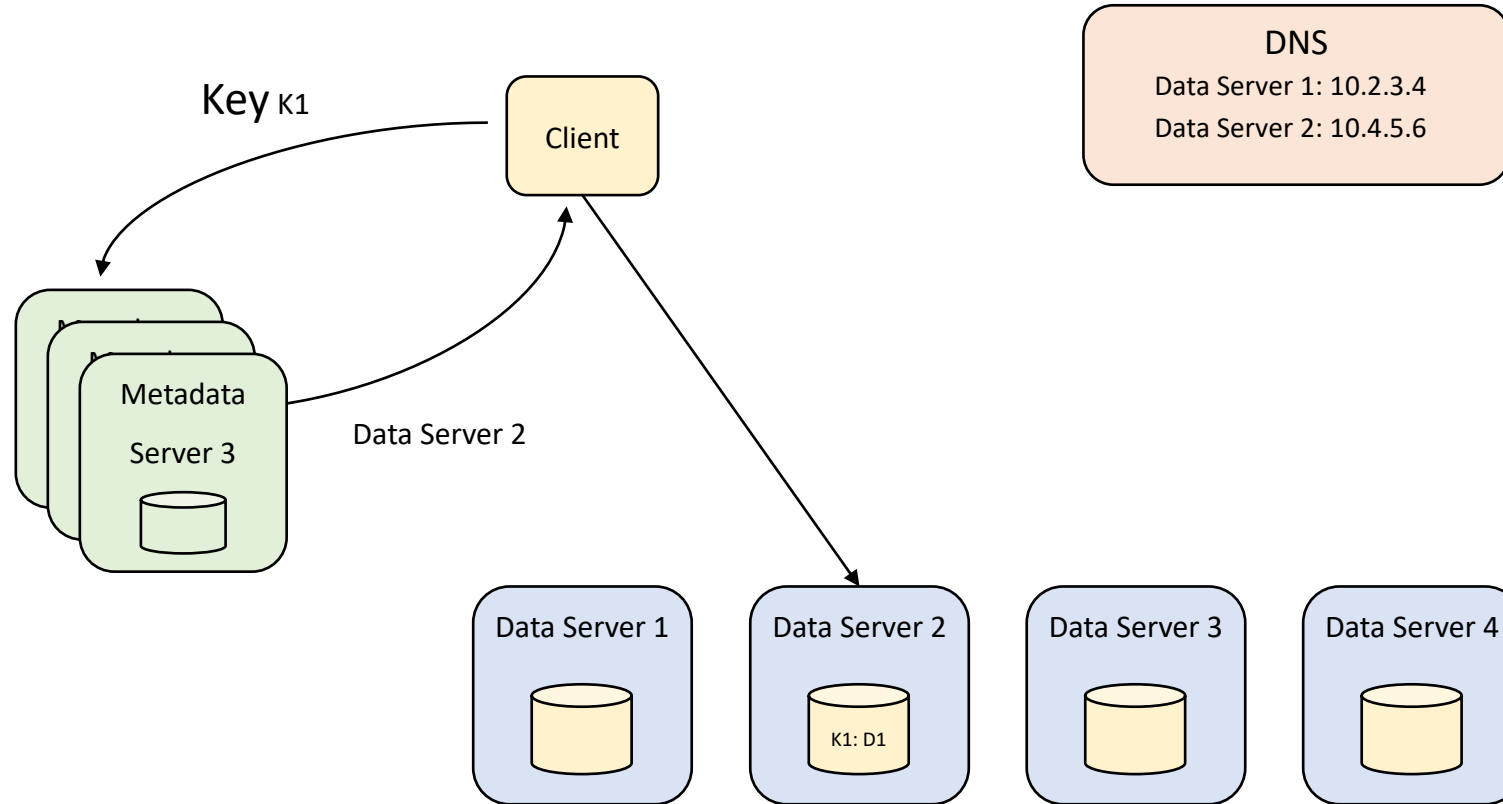
Hadoop/HBase Stateful Applications



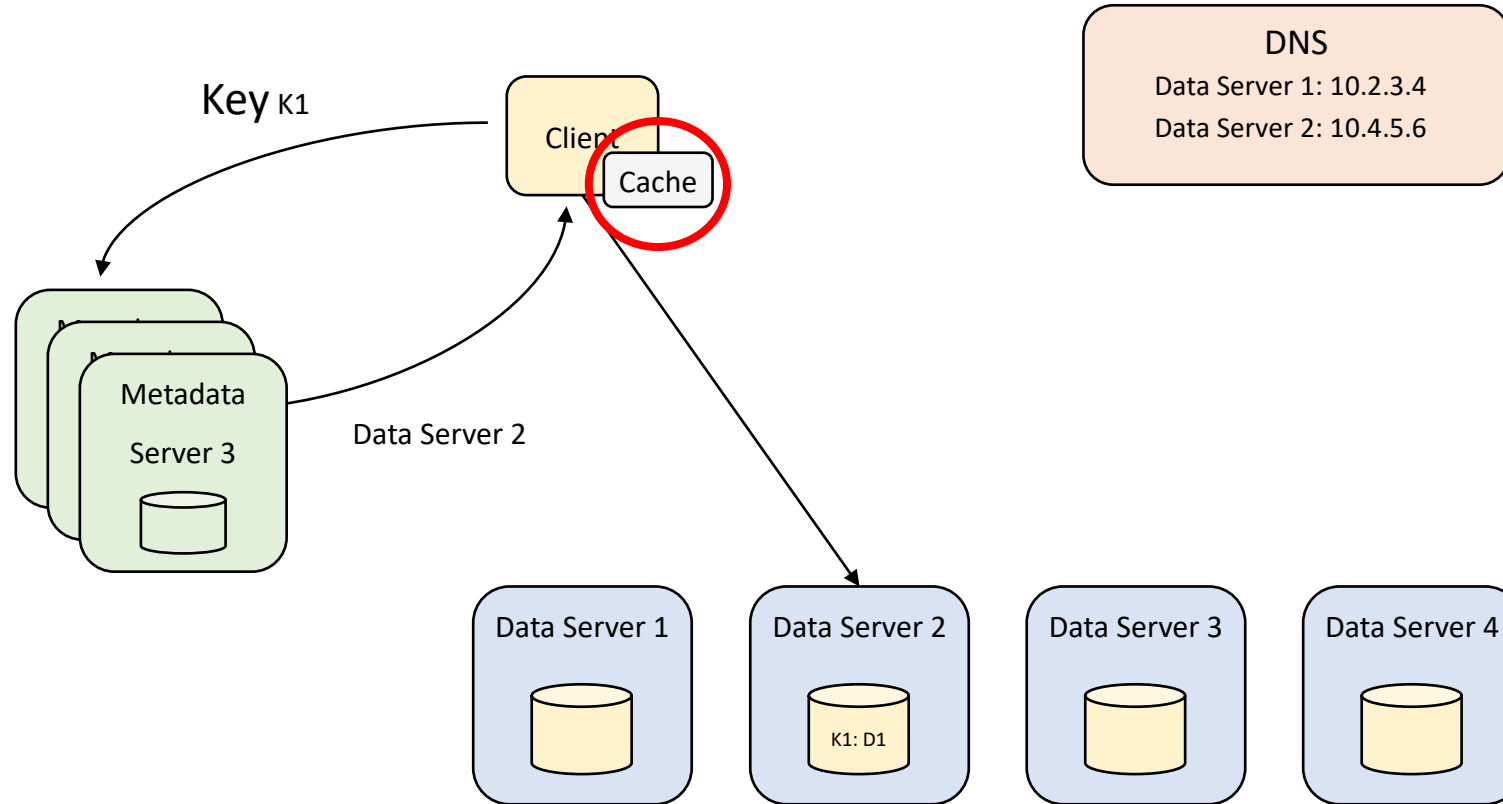
Hadoop/HBase Stateful Applications



Hadoop/HBase Stateful Applications



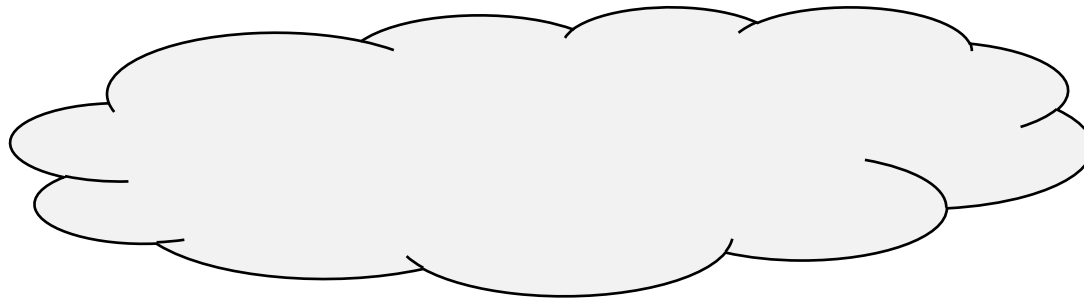
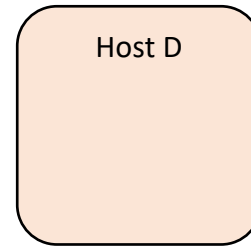
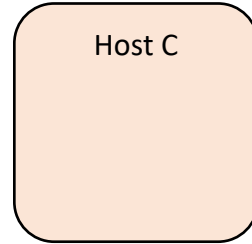
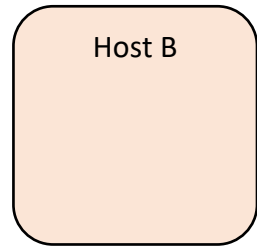
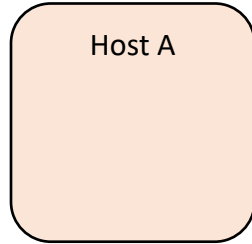
Hadoop/HBase Stateful Applications



Kubernetes StatefulSets



Replicas: 3 Pods
Disk: 1TB of SSD

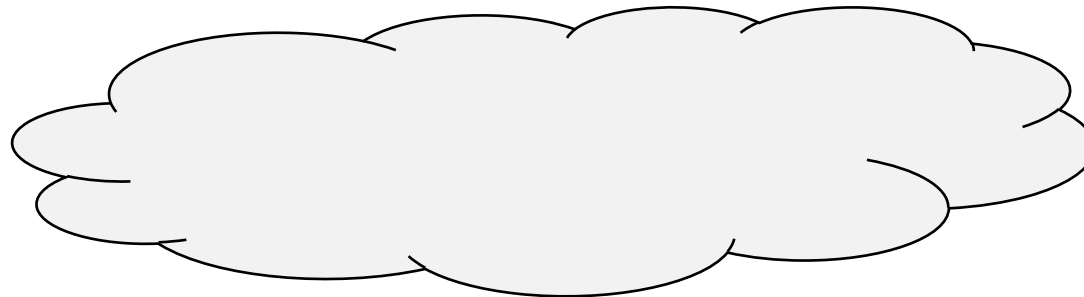
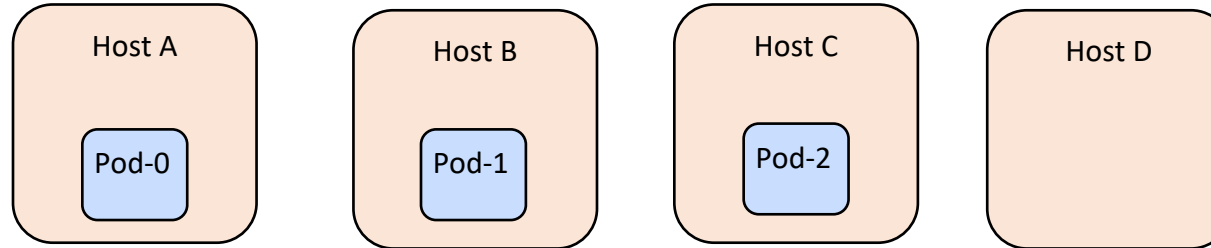


Kubernetes StatefulSets

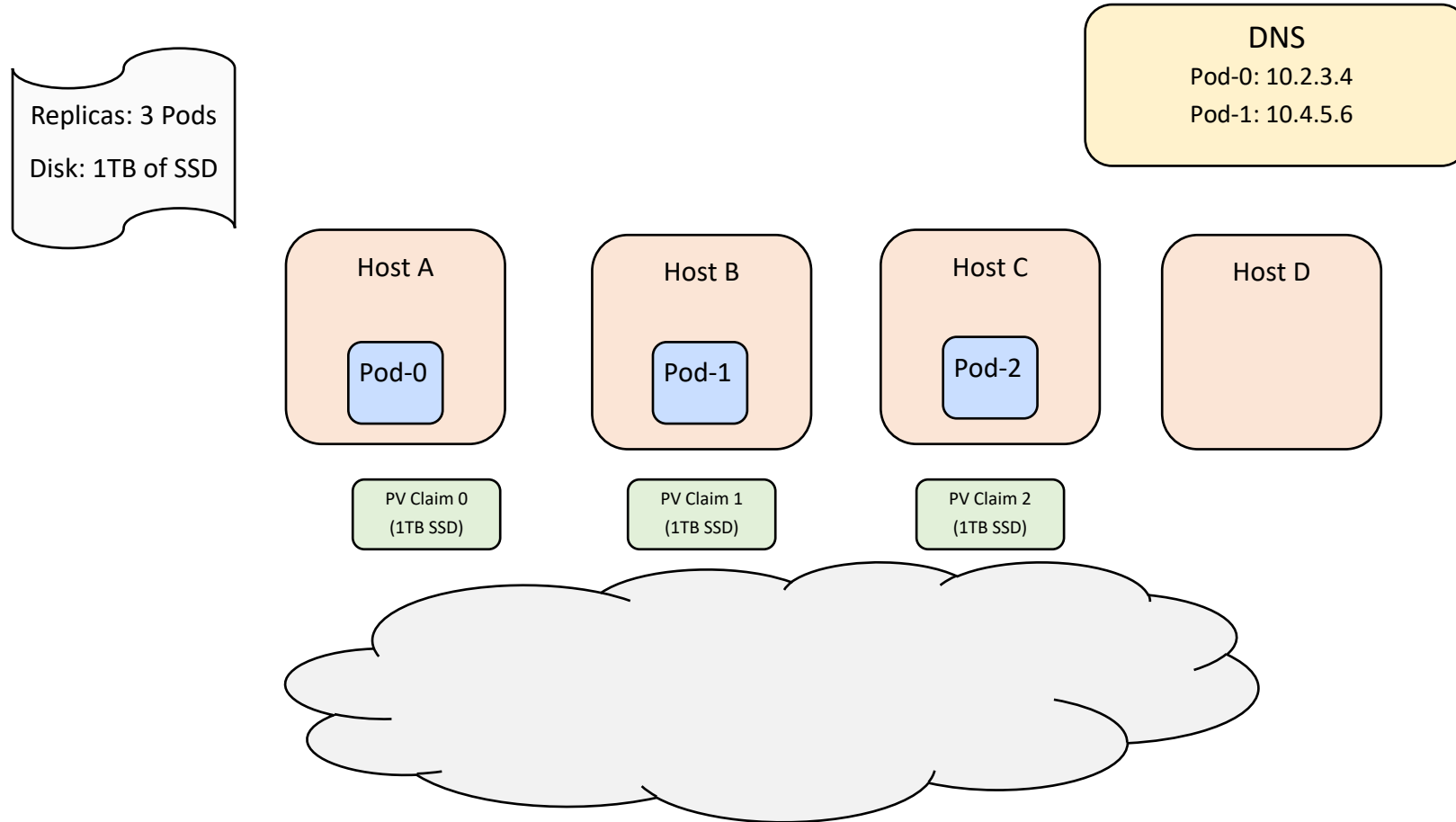


Replicas: 3 Pods
Disk: 1TB of SSD

DNS
Pod-0: 10.2.3.4
Pod-1: 10.4.5.6



Kubernetes StatefulSets

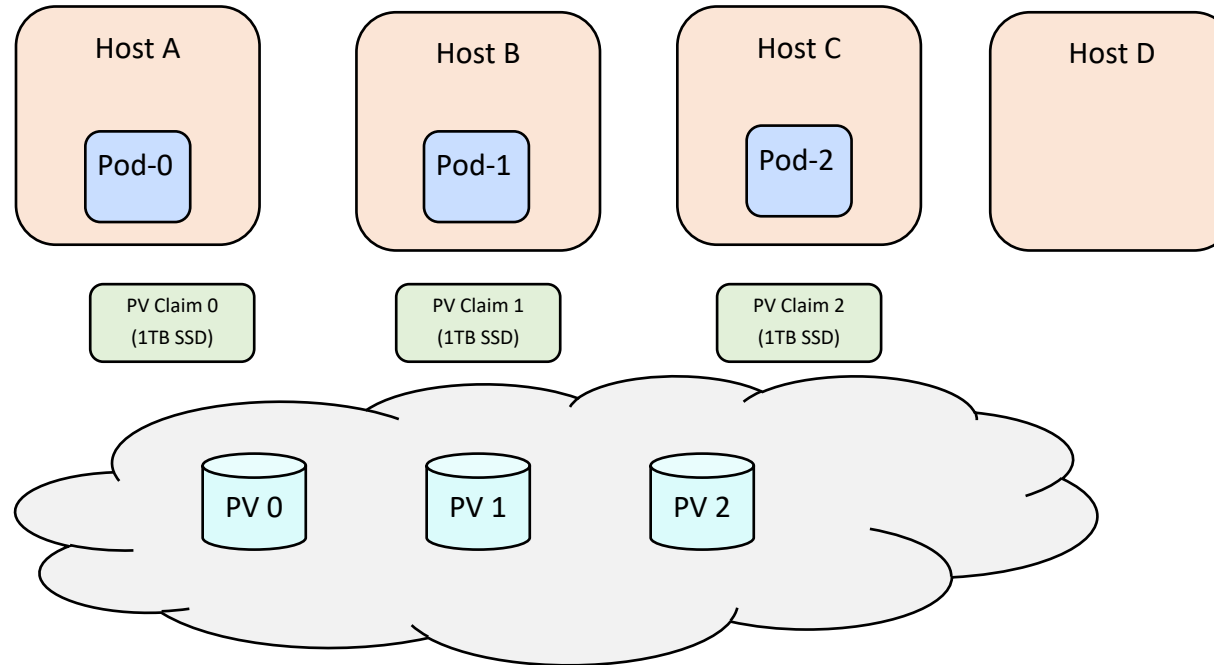


Kubernetes StatefulSets



Replicas: 3 Pods
Disk: 1TB of SSD

DNS
Pod-0: 10.2.3.4
Pod-1: 10.4.5.6

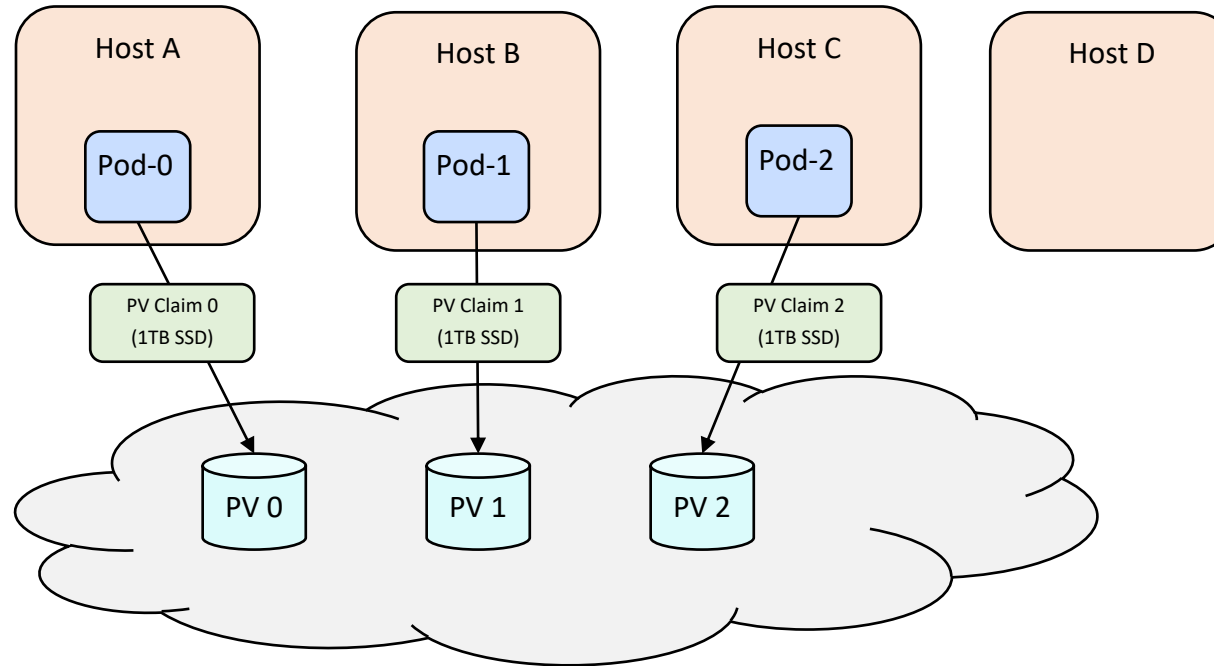


Kubernetes StatefulSets

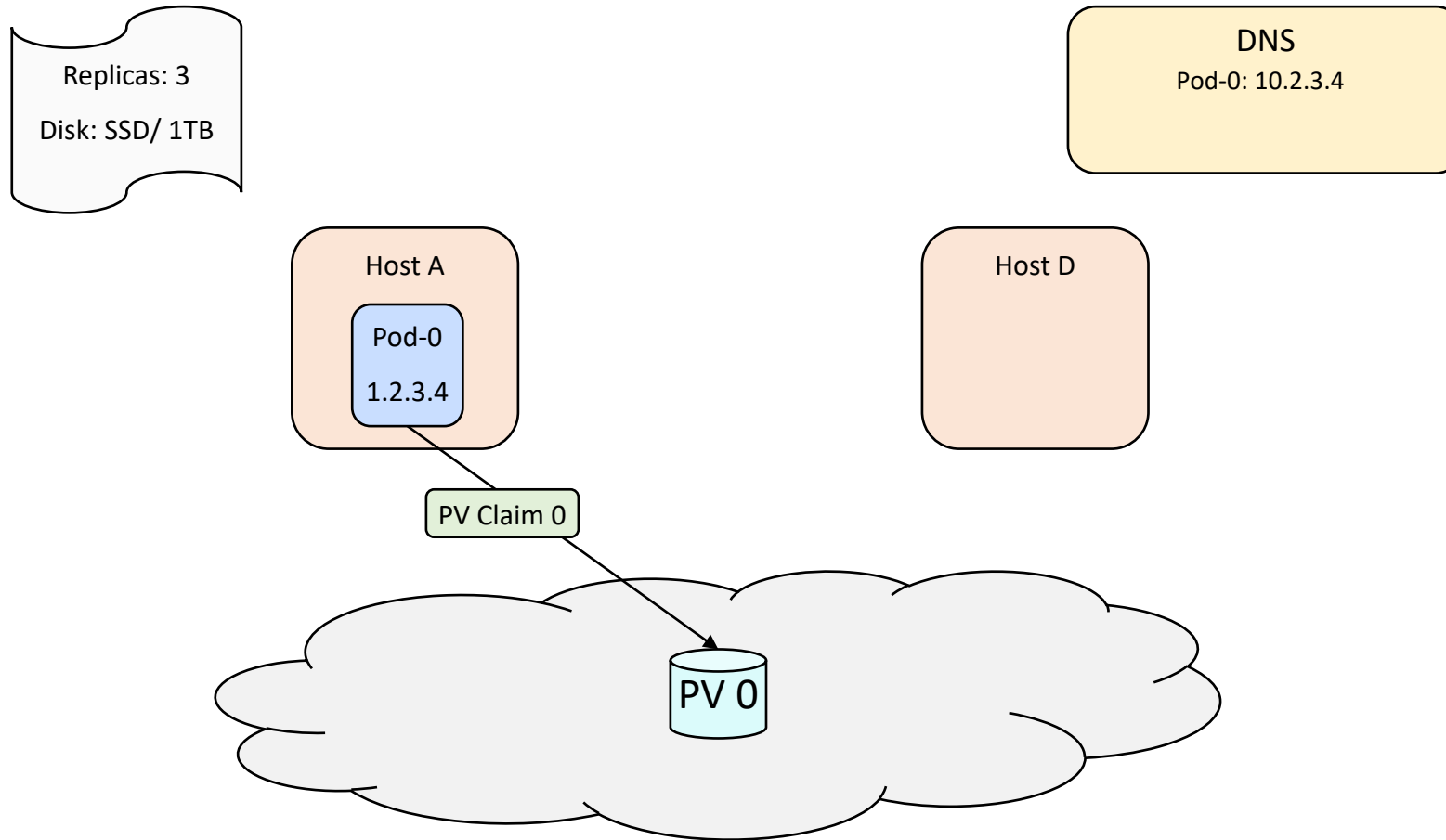


Replicas: 3 Pods
Disk: 1TB of SSD

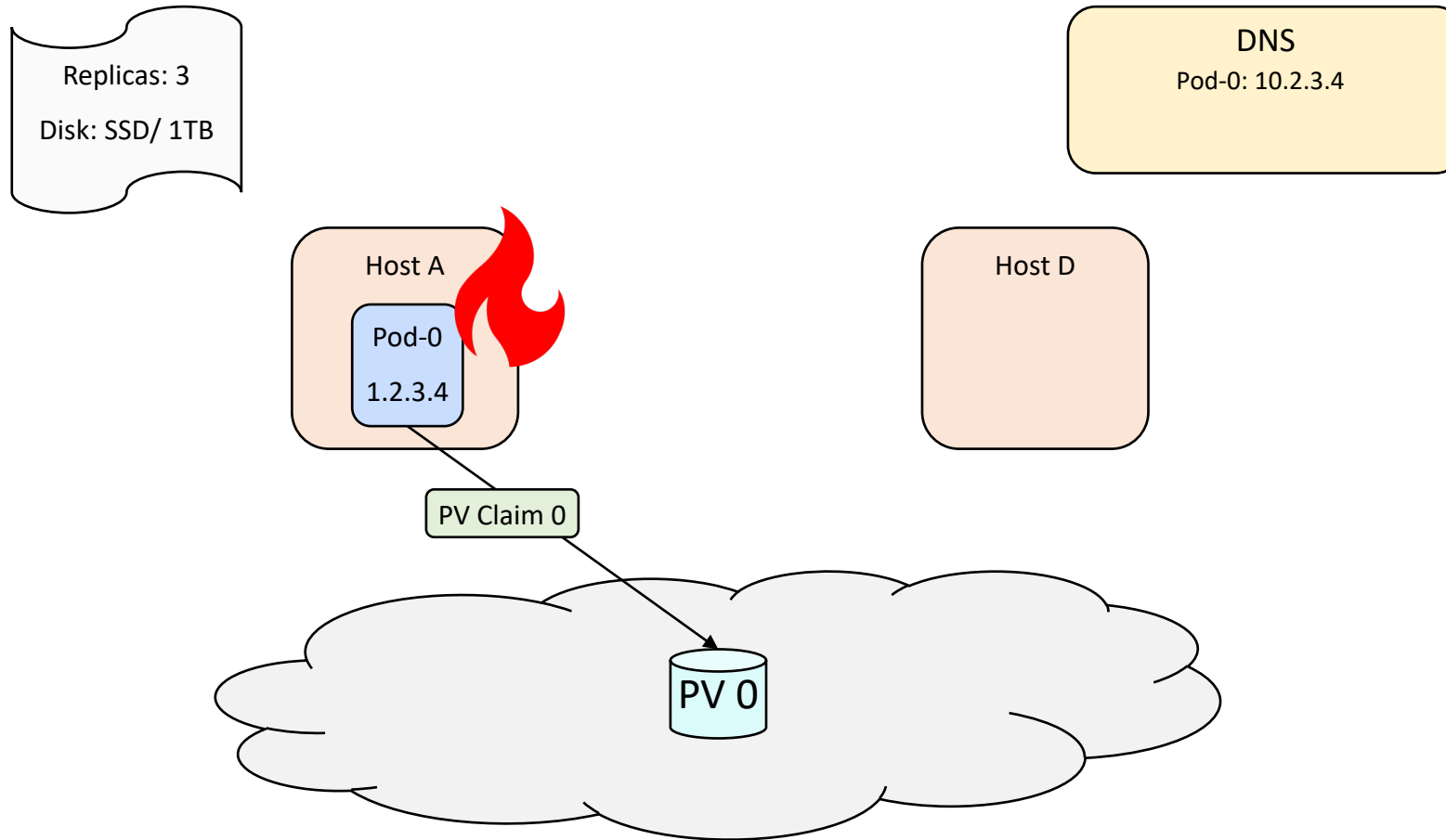
DNS
Pod-0: 10.2.3.4
Pod-1: 10.4.5.6



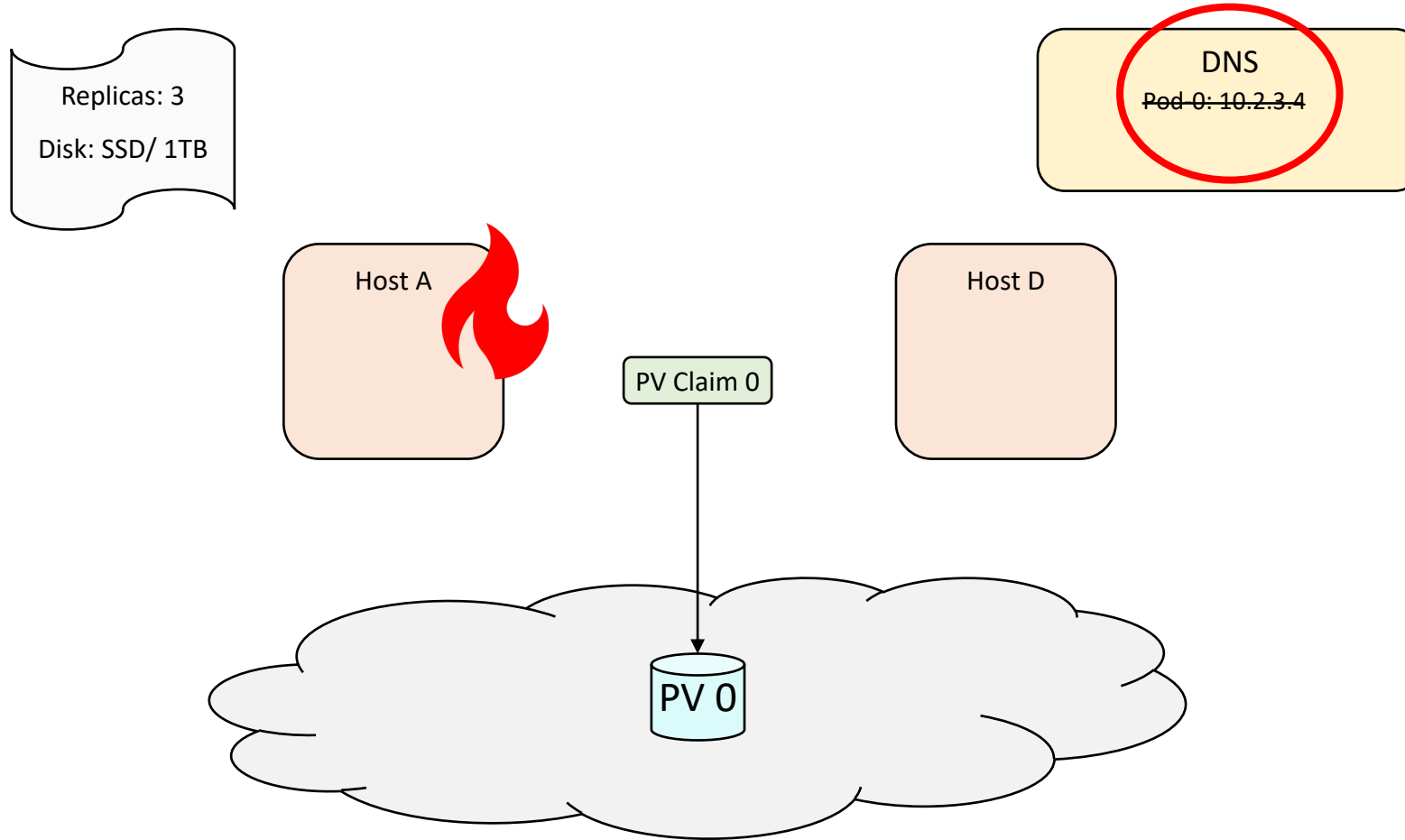
Kubernetes StatefulSets (STS)



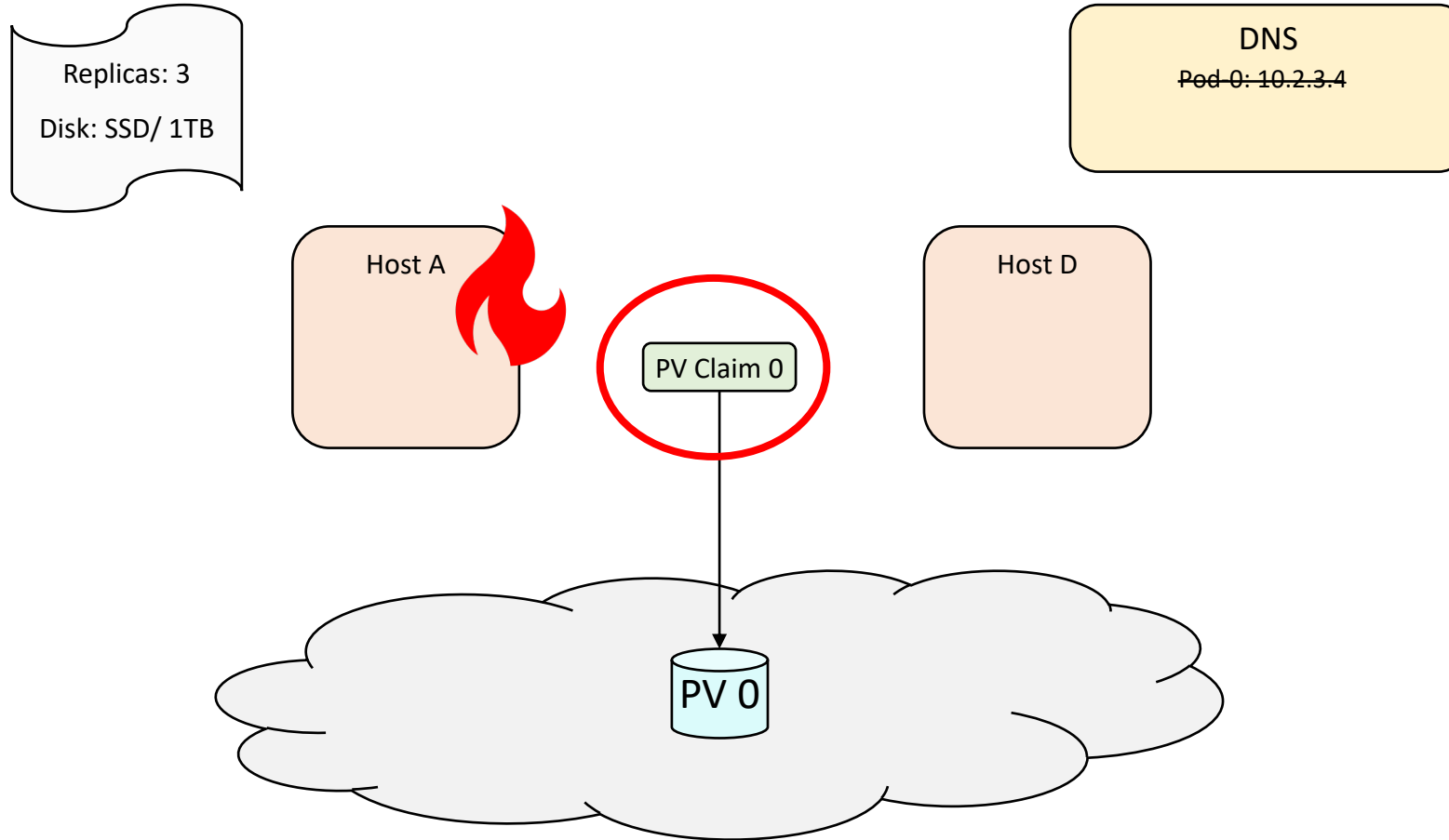
Kubernetes StatefulSets (STS)



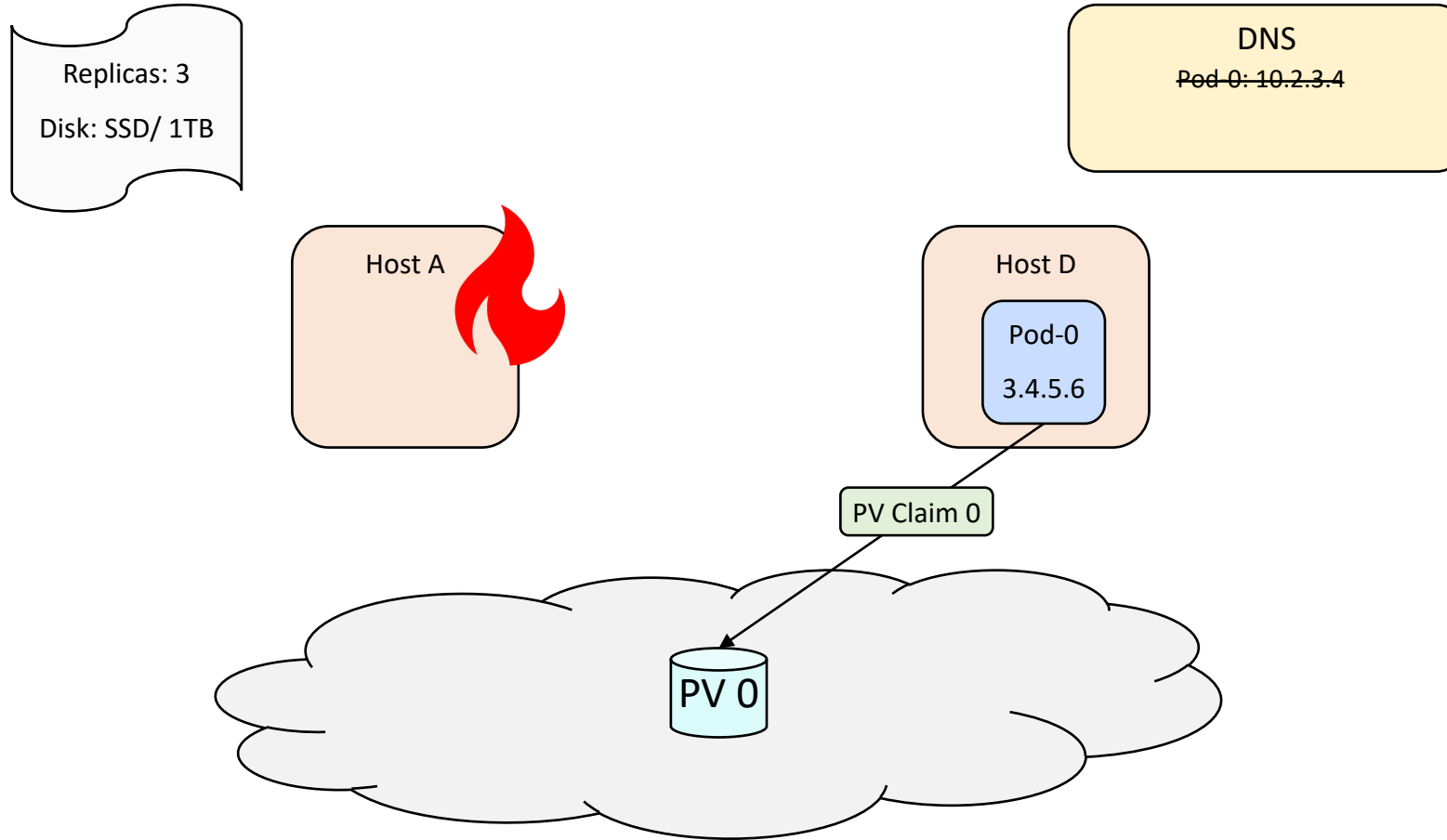
Kubernetes StatefulSets (STS)



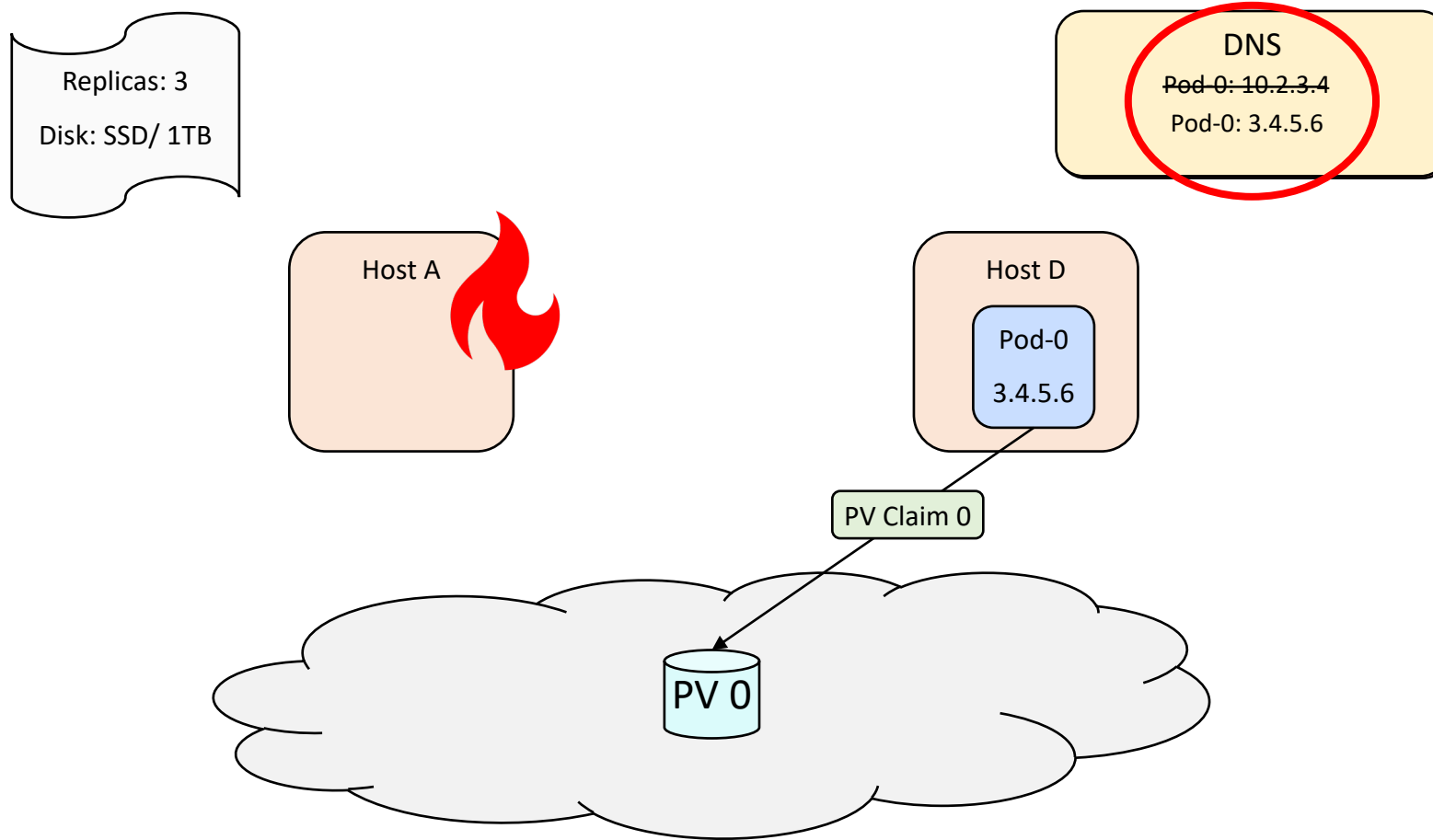
Kubernetes StatefulSets (STS)



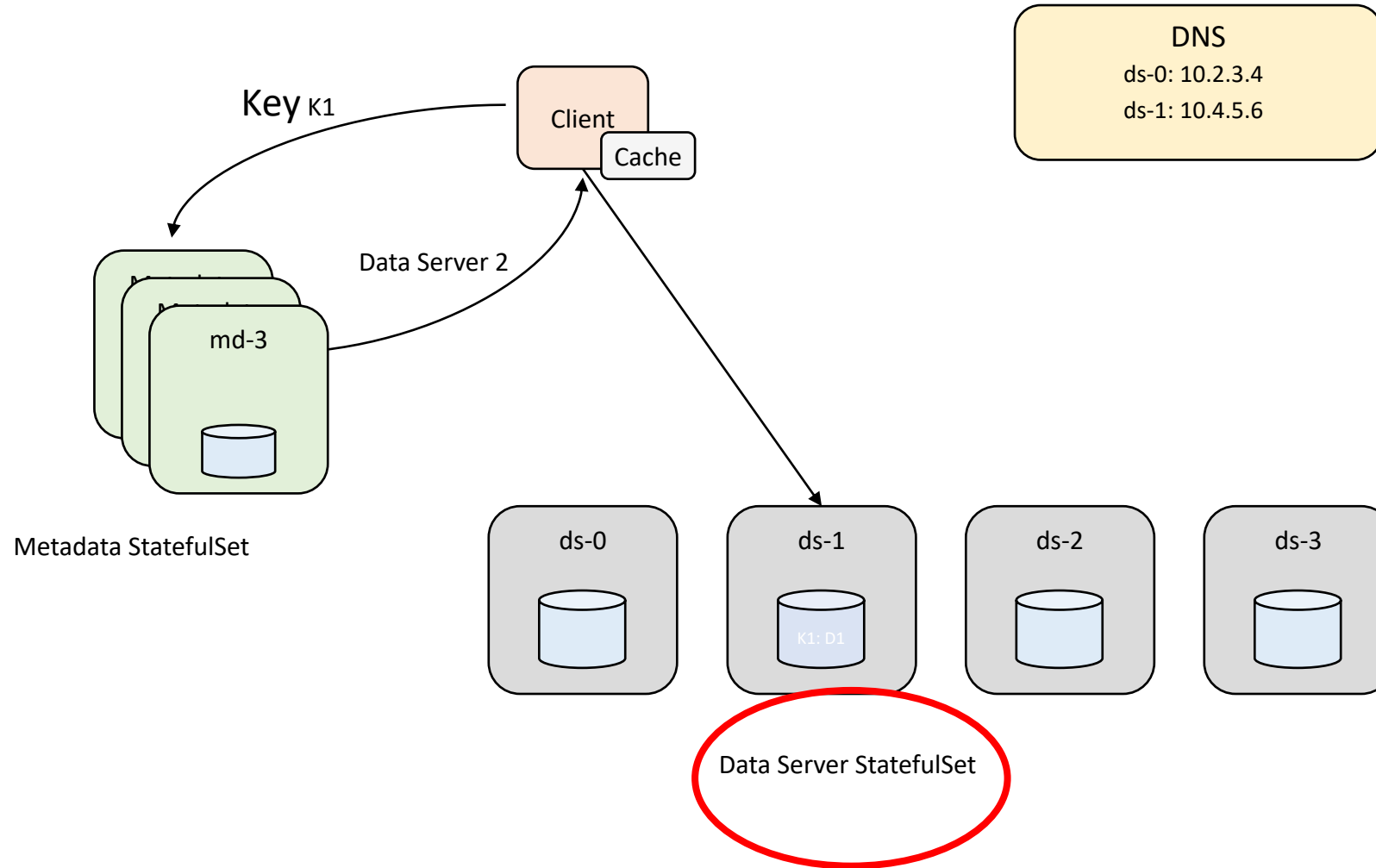
Kubernetes StatefulSets (STS)



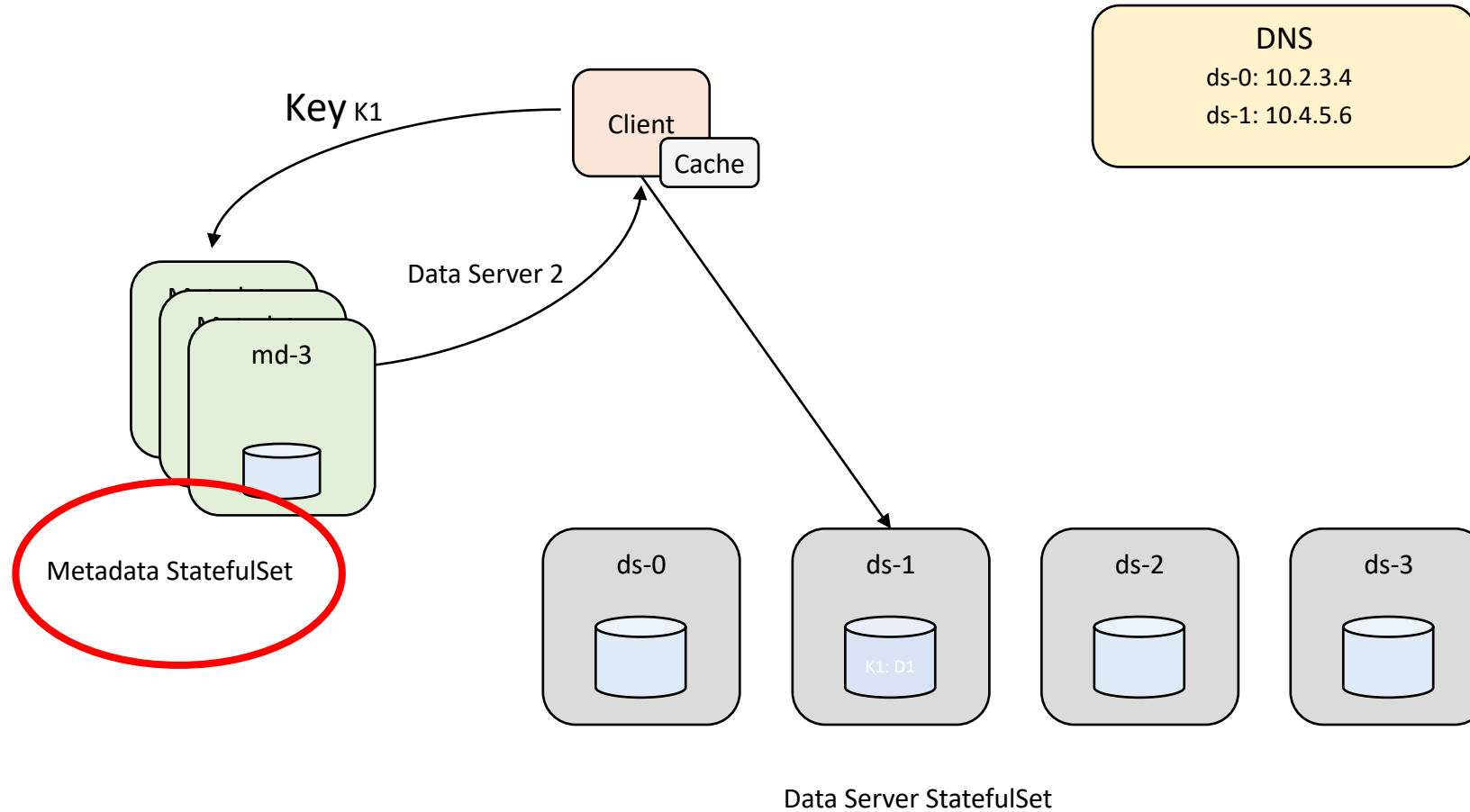
Kubernetes StatefulSets (STS)



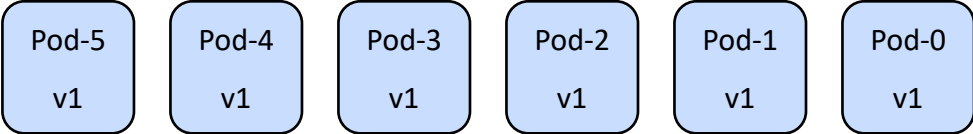
Hadoop/HBase Stateful Applications



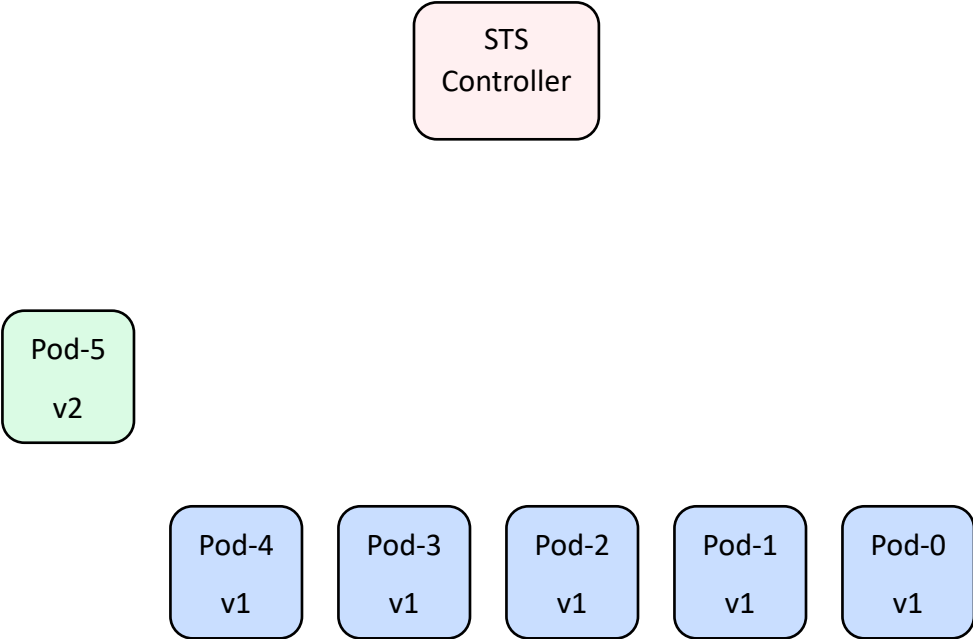
Hadoop/HBase Stateful Applications



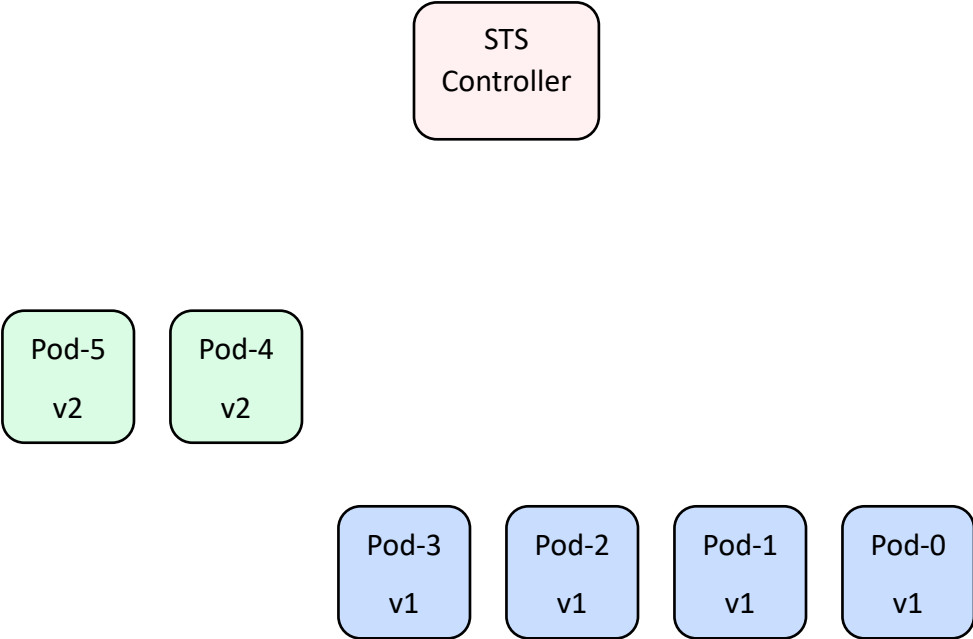
StatefulSet (STS) Controller – Slow Upgrade



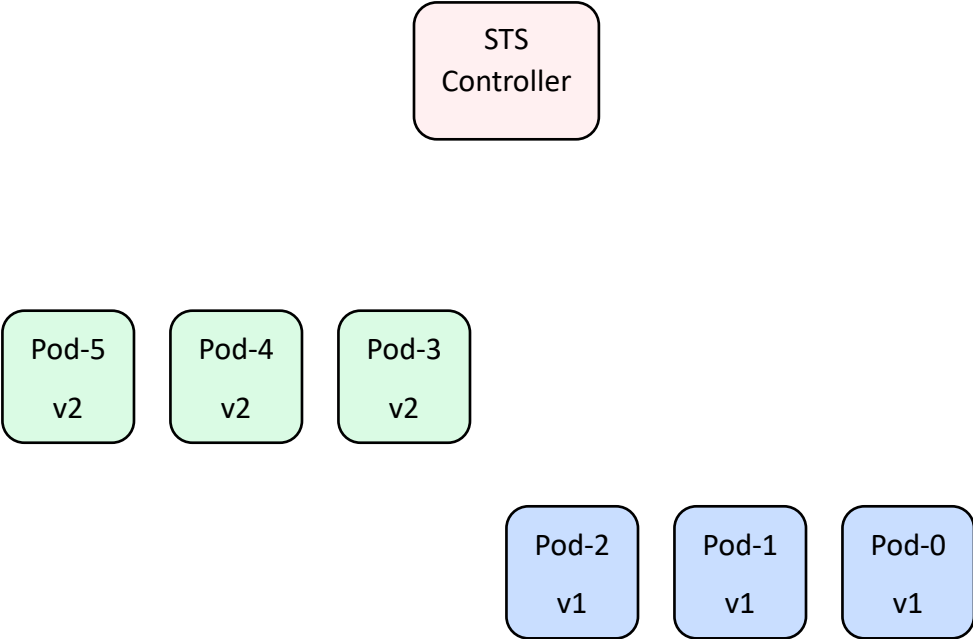
StatefulSet (STS) Controller – Slow Upgrade



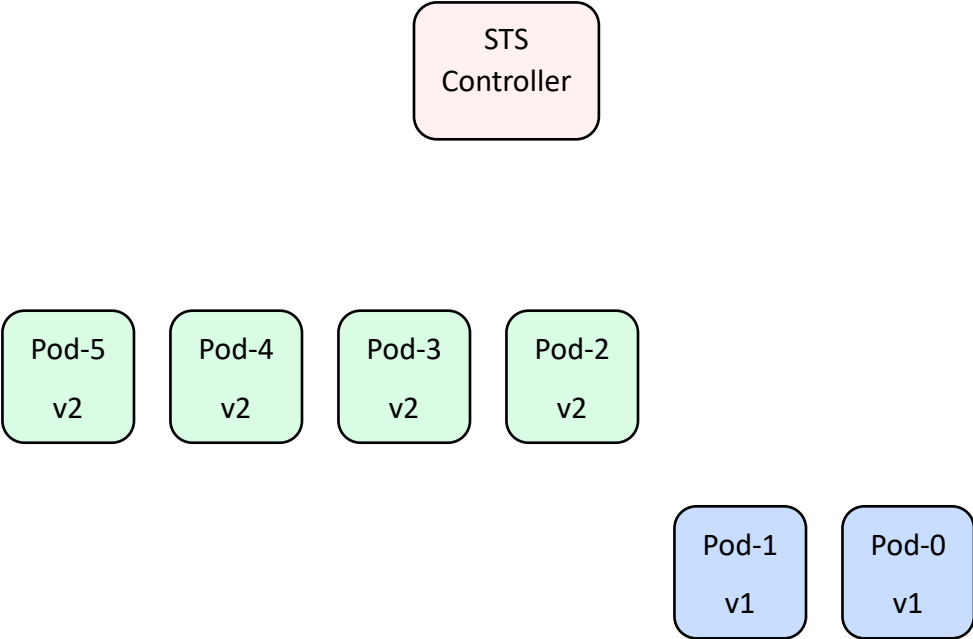
StatefulSet (STS) Controller – Slow Upgrade



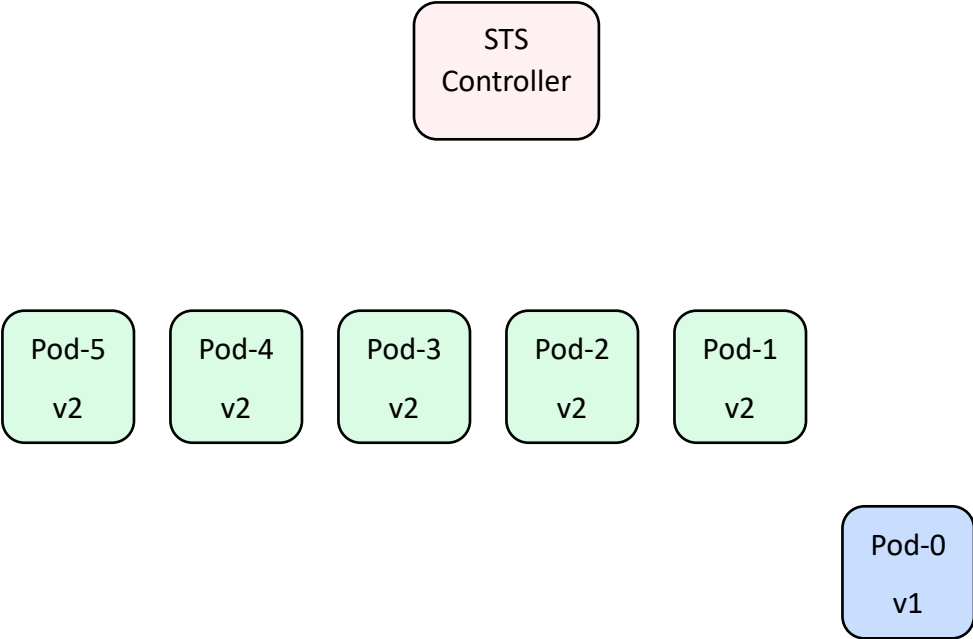
StatefulSet (STS) Controller – Slow Upgrade



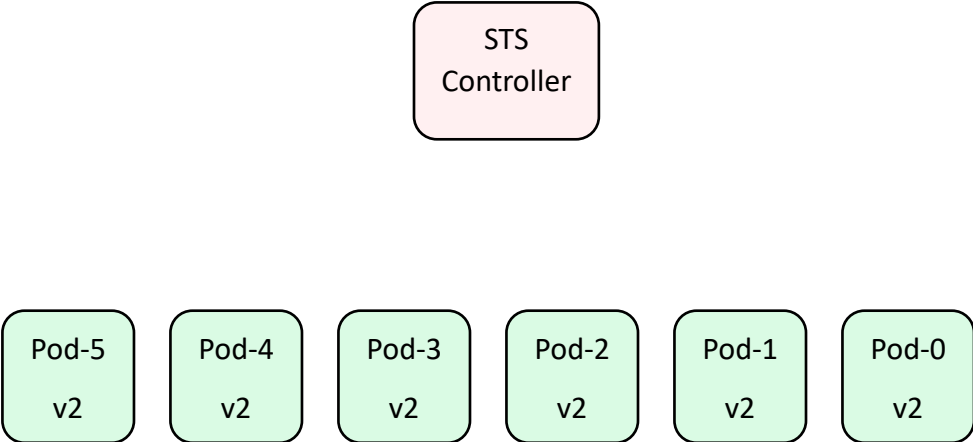
StatefulSet (STS) Controller – Slow Upgrade



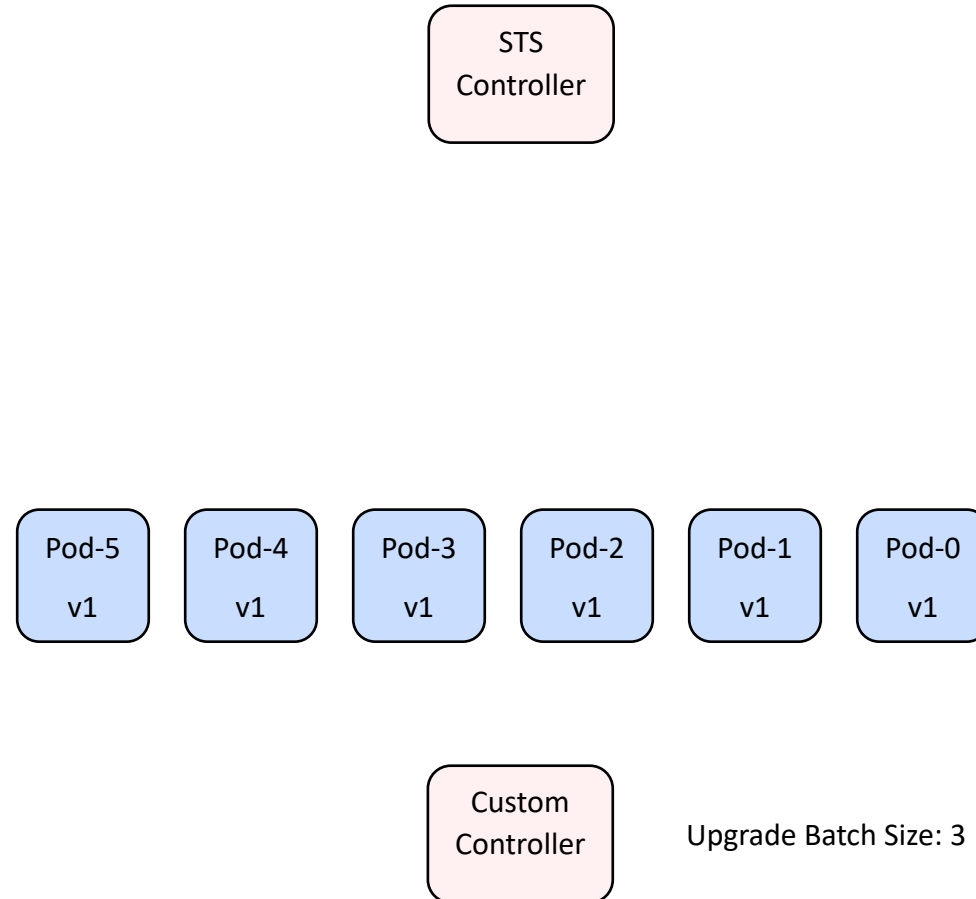
StatefulSet (STS) Controller – Slow Upgrade



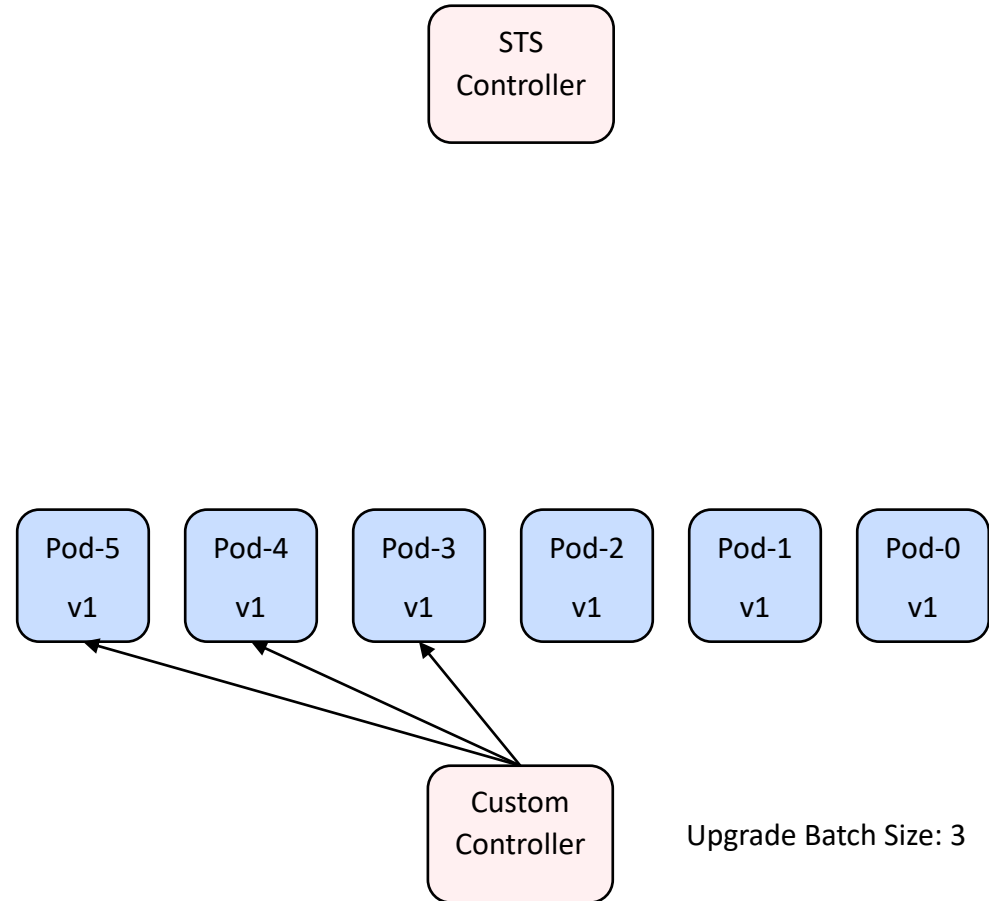
StatefulSet (STS) Controller – Slow Upgrade



Custom Controller - Batching



Custom Controller - Batching



Custom Controller - Batching



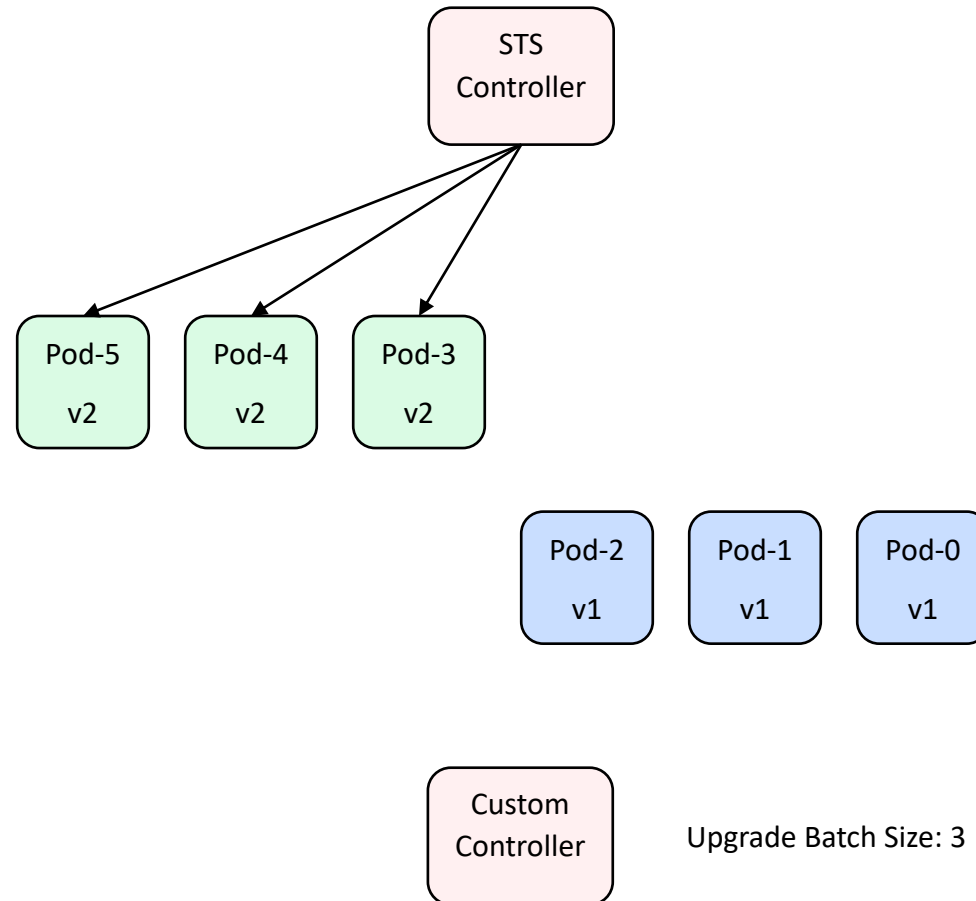
STS
Controller



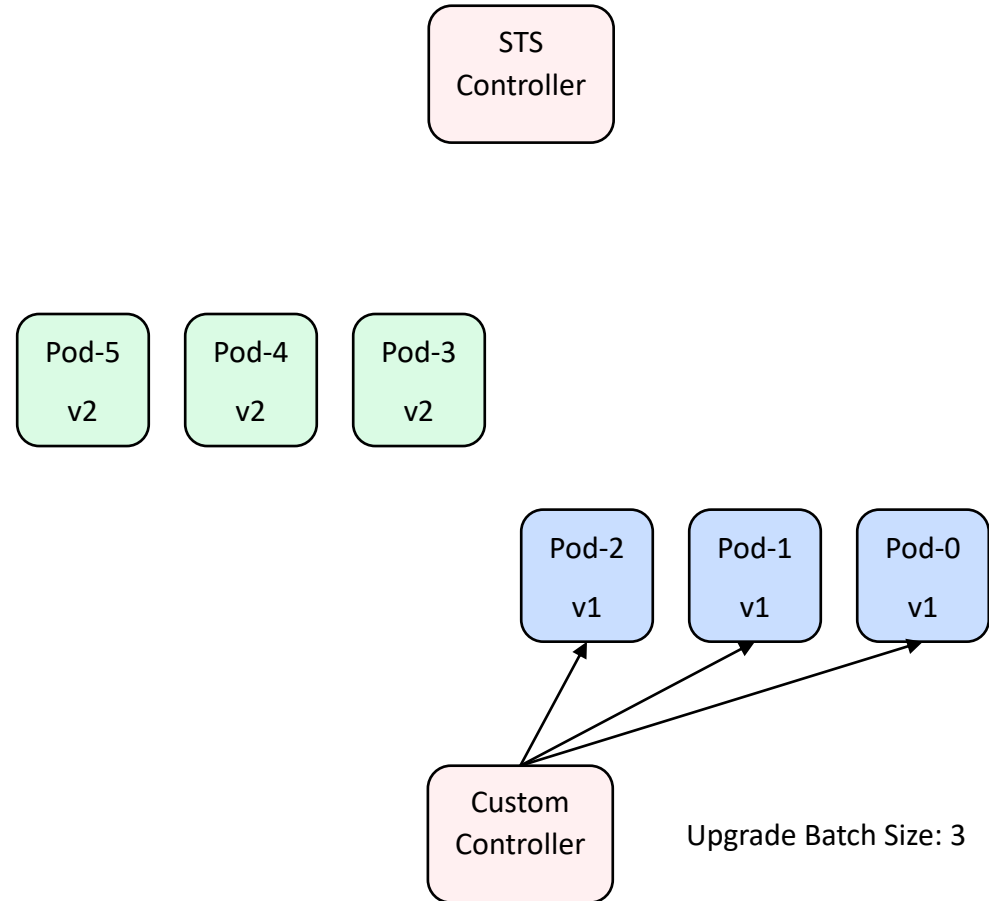
Custom
Controller

Upgrade Batch Size: 3

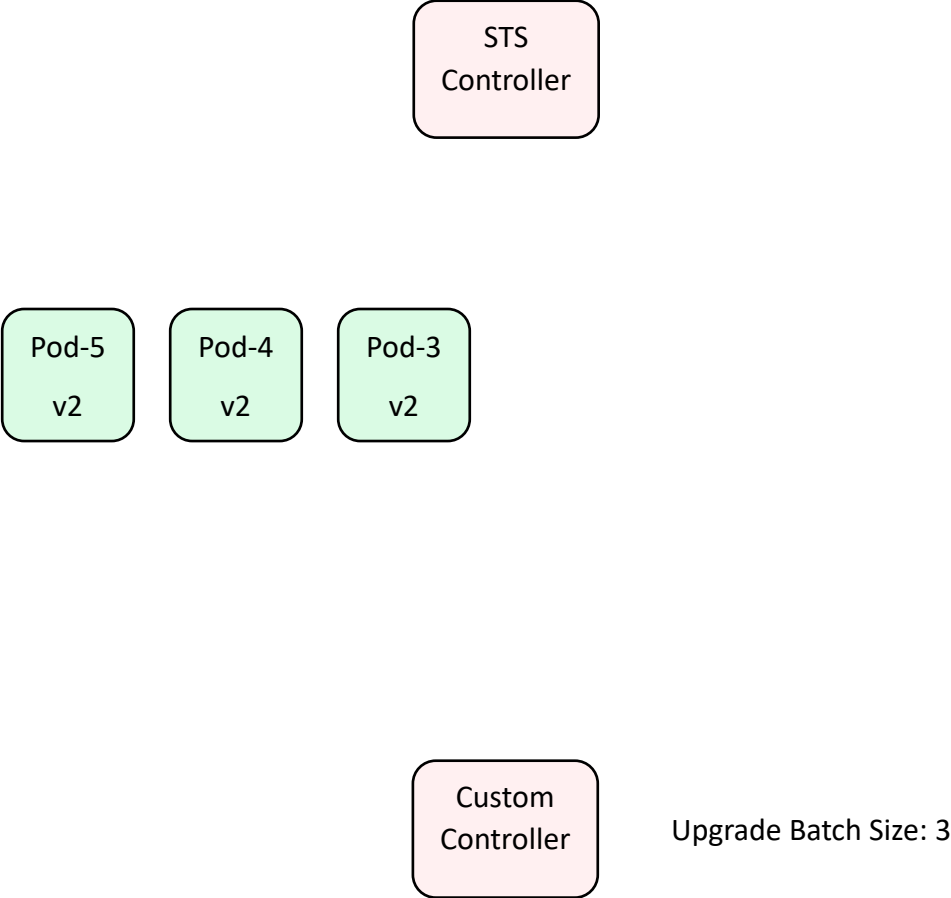
Custom Controller - Batching



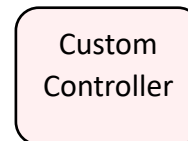
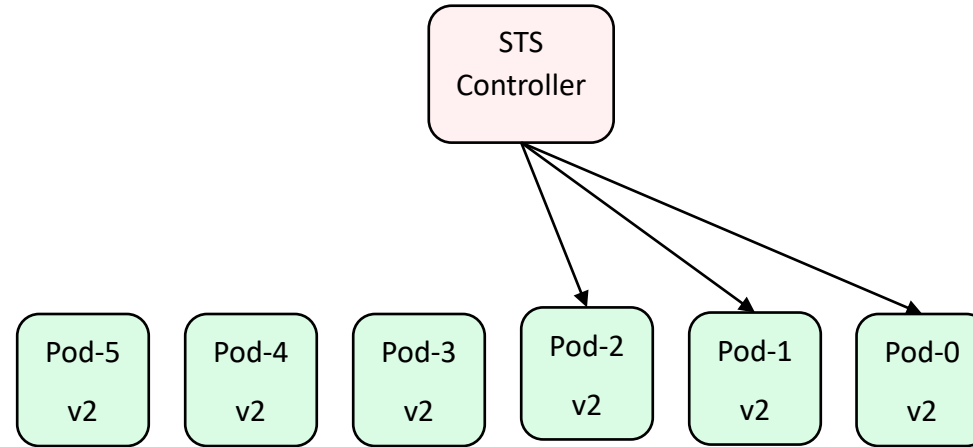
Custom Controller - Batching



Custom Controller - Batching



Custom Controller - Batching

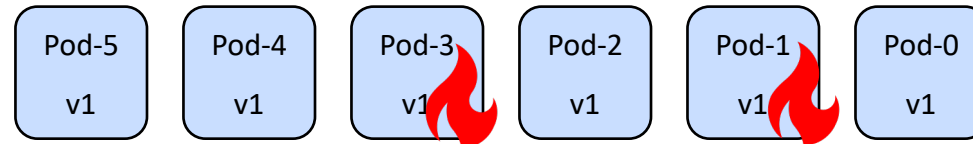


Upgrade Batch Size: 3

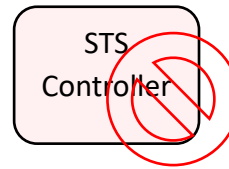
STS Controller – Pod Disruption Budget



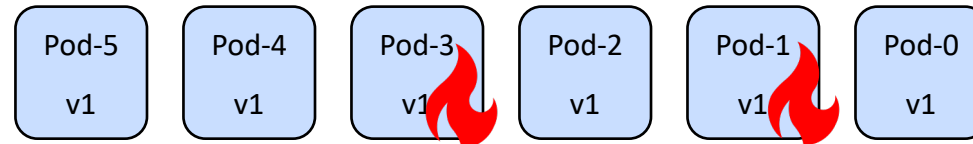
Disruption Budget: 1



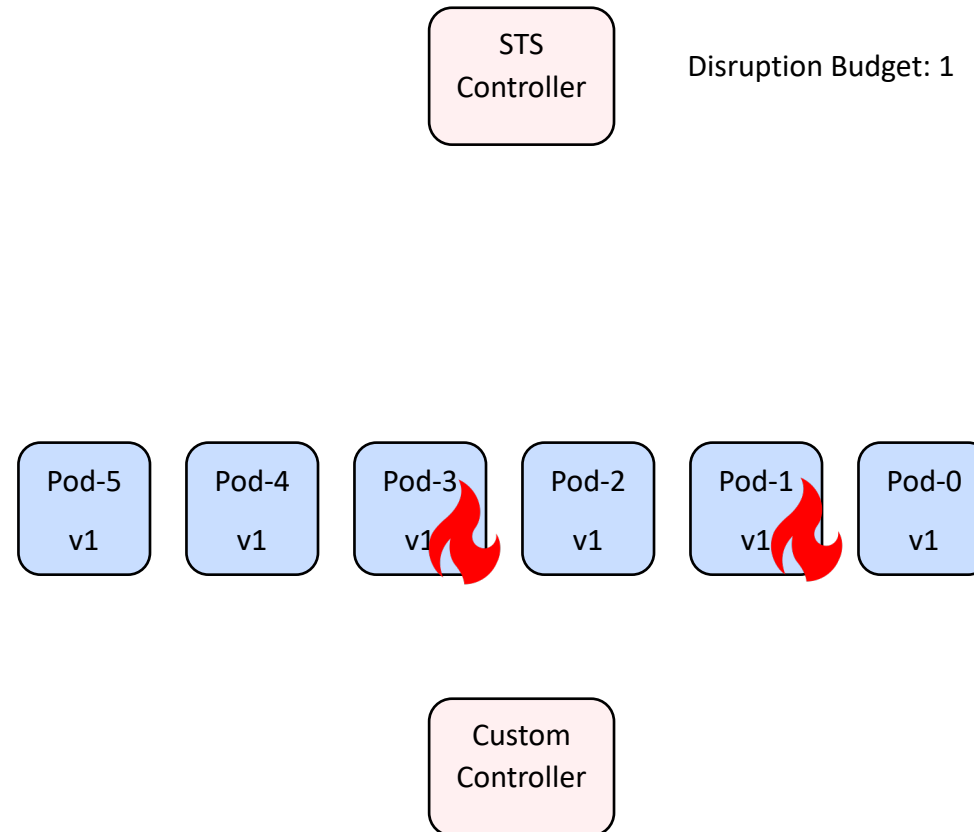
STS Controller – Pod Disruption Budget



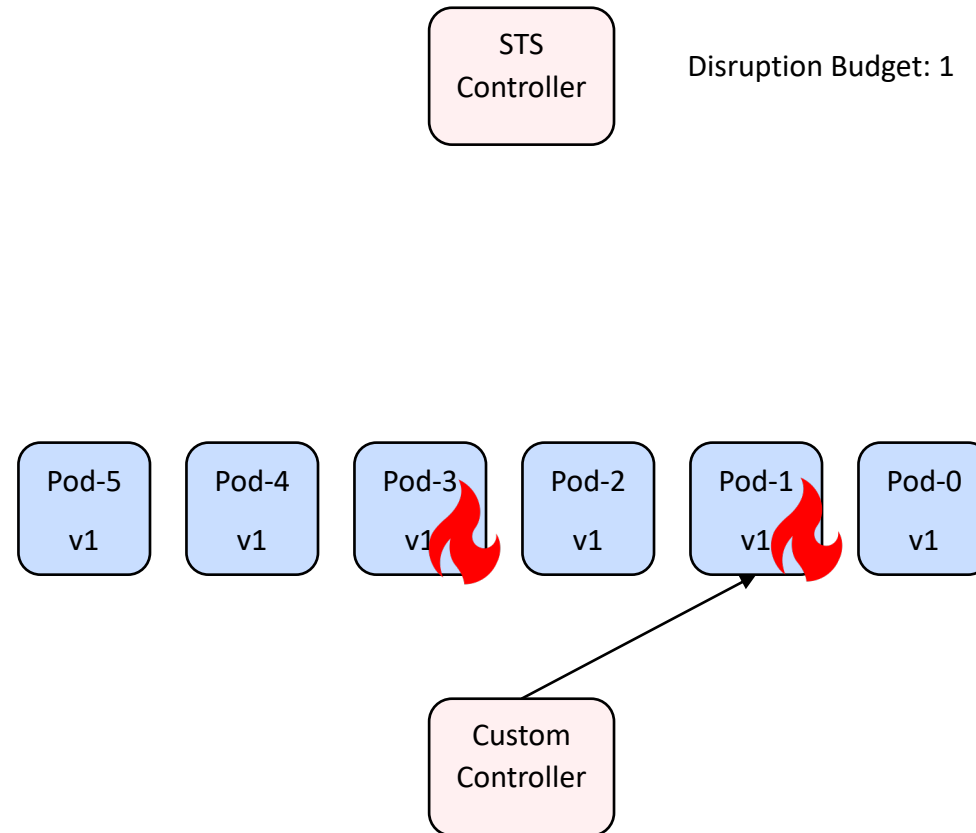
Disruption Budget: 1



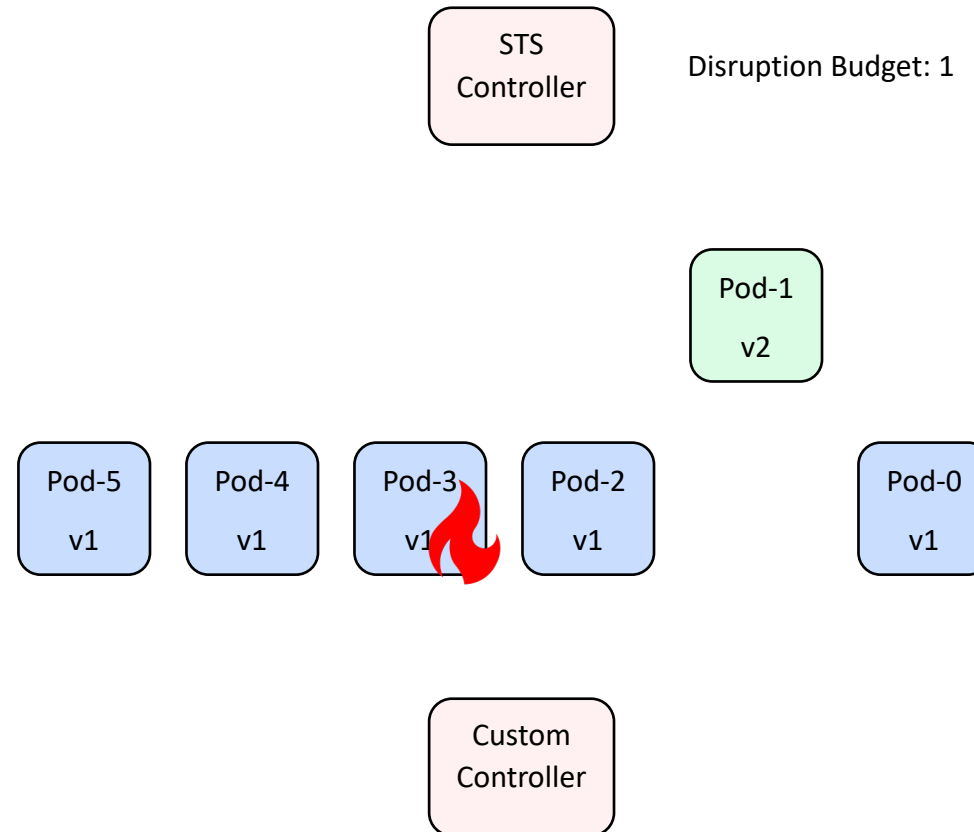
STS Controller – Pod Disruption Budget



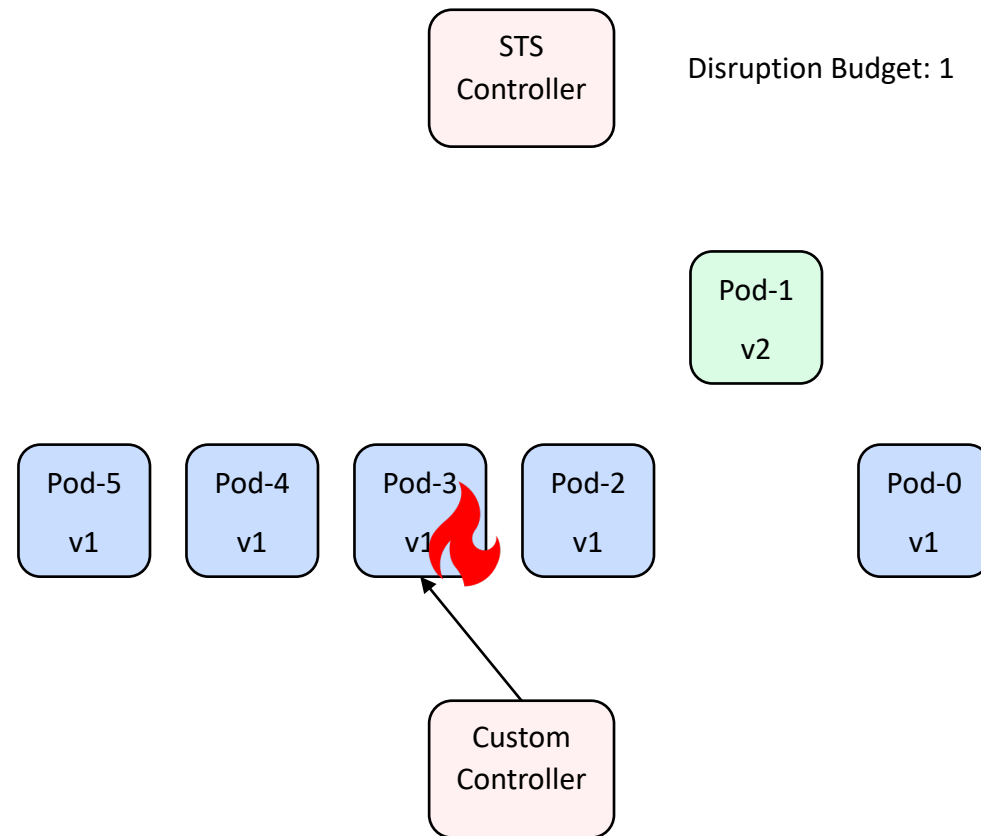
STS Controller – Pod Disruption Budget



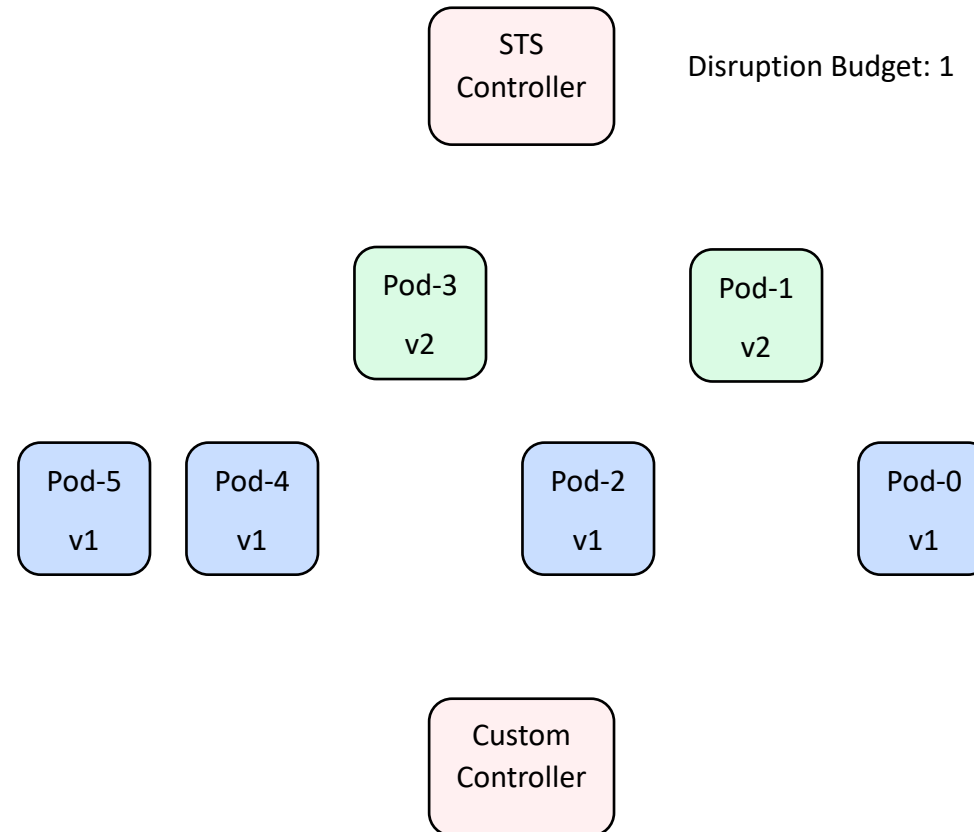
STS Controller – Pod Disruption Budget



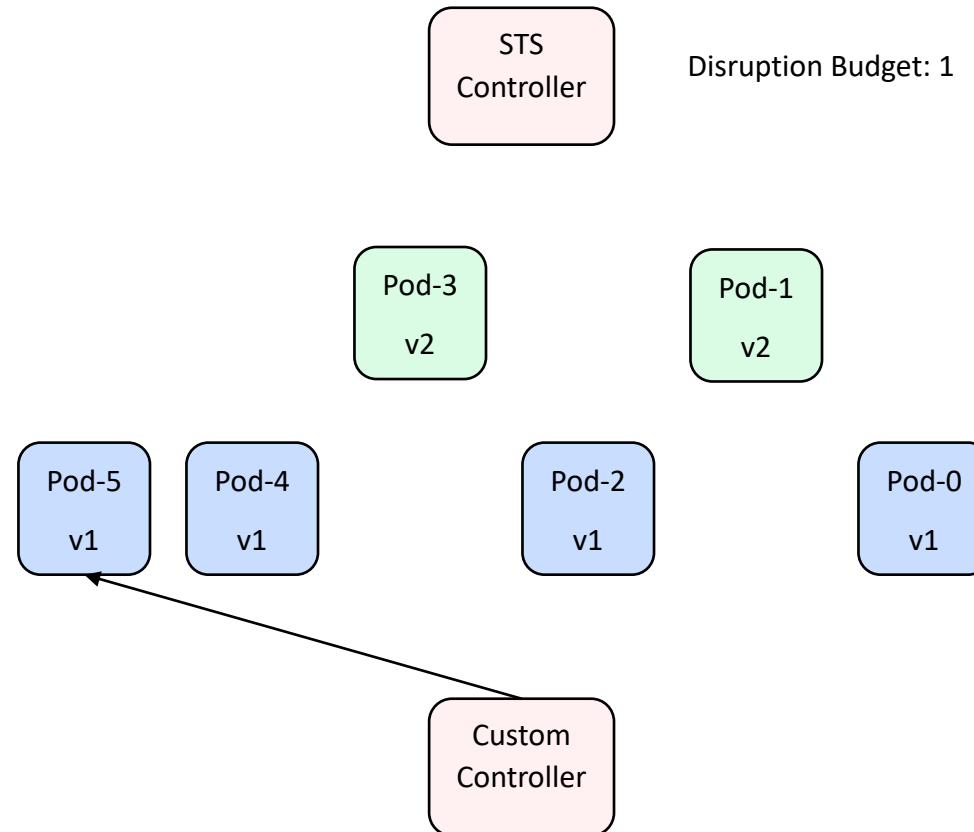
STS Controller – Pod Disruption Budget



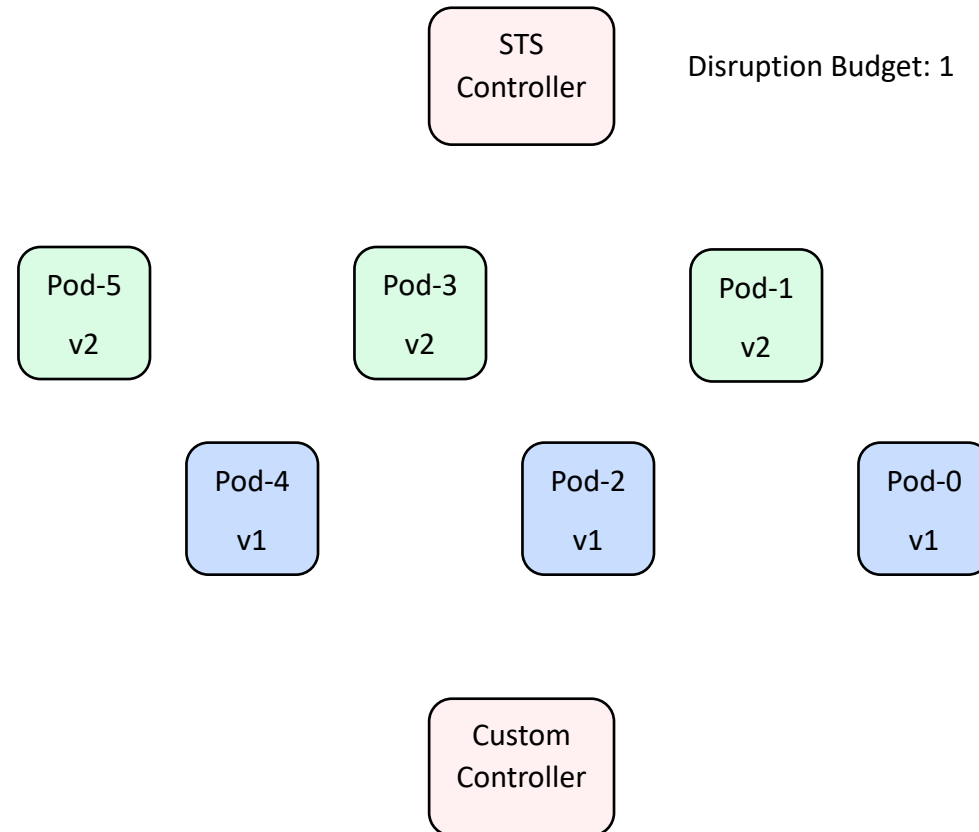
STS Controller – Pod Disruption Budget



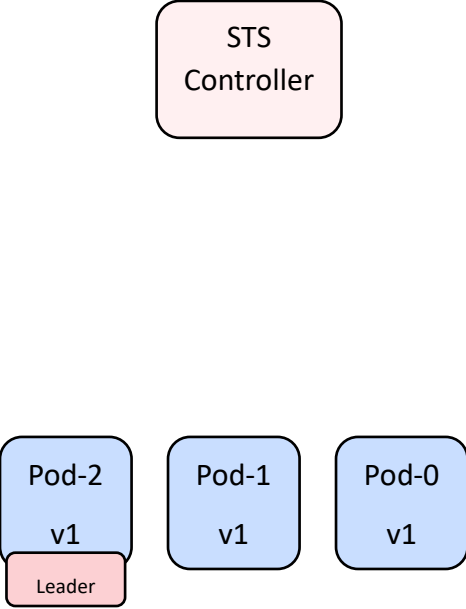
STS Controller – Pod Disruption Budget



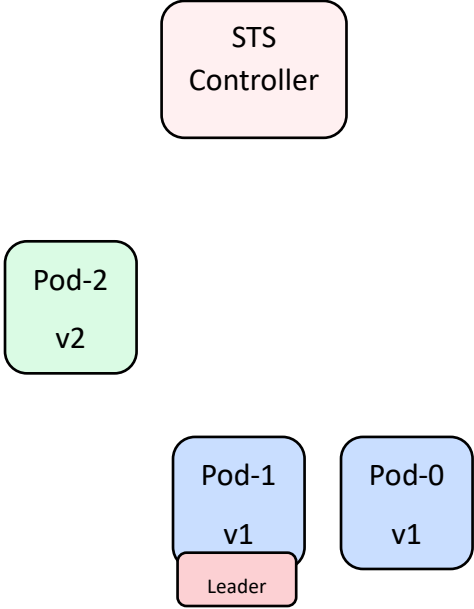
STS Controller – Pod Disruption Budget



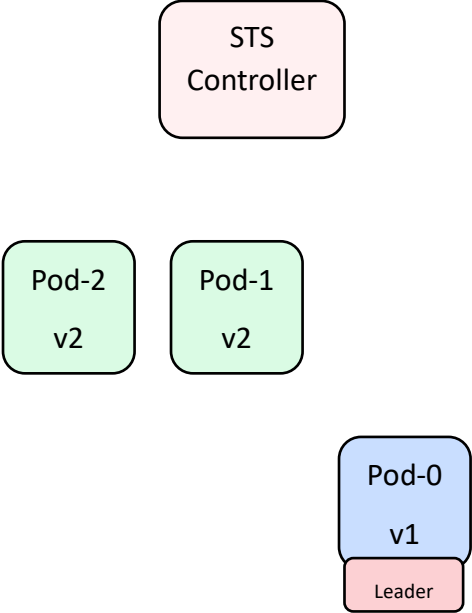
Custom Controller – Leader Problem



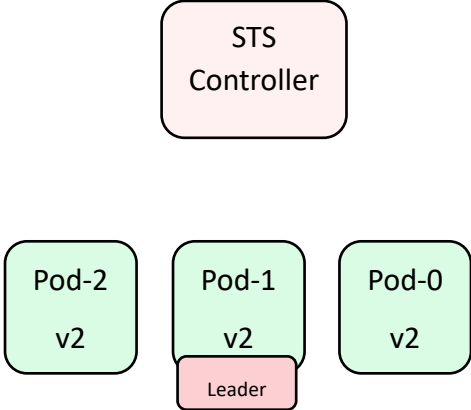
Custom Controller – Leader Problem



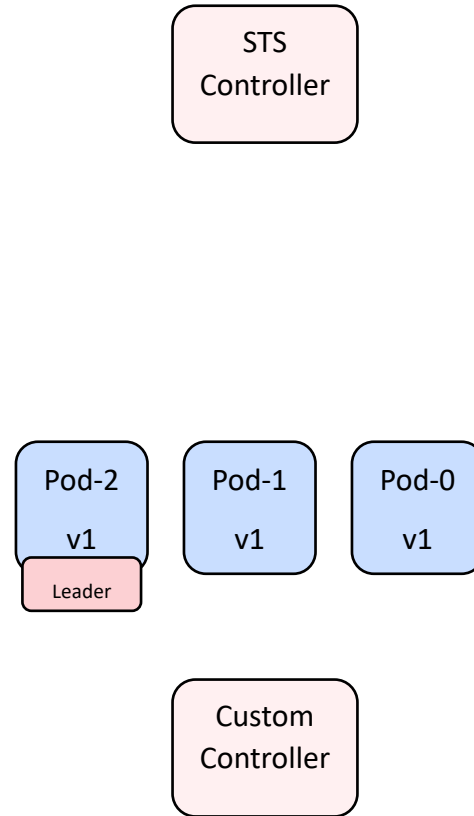
Custom Controller – Leader Problem



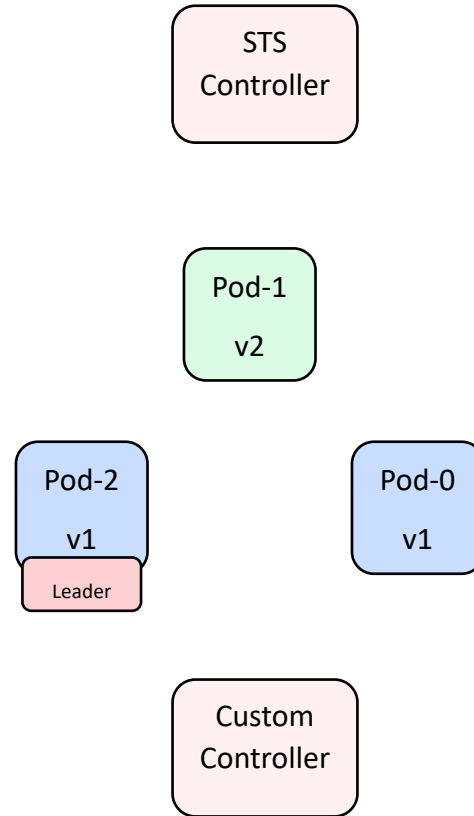
Custom Controller – Leader Problem



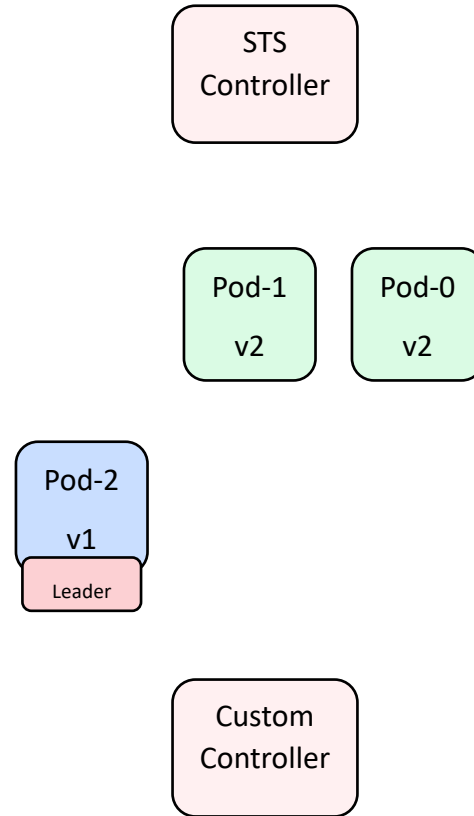
Custom Controller – Leader Solution



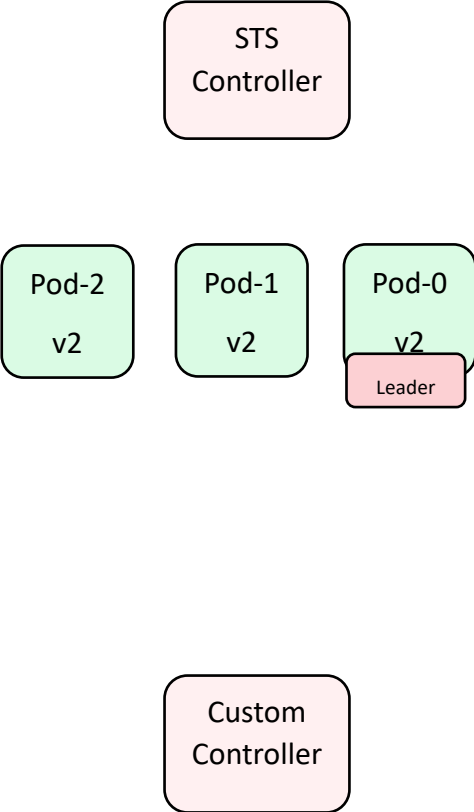
Custom Controller – Leader Solution



Custom Controller – Leader Solution



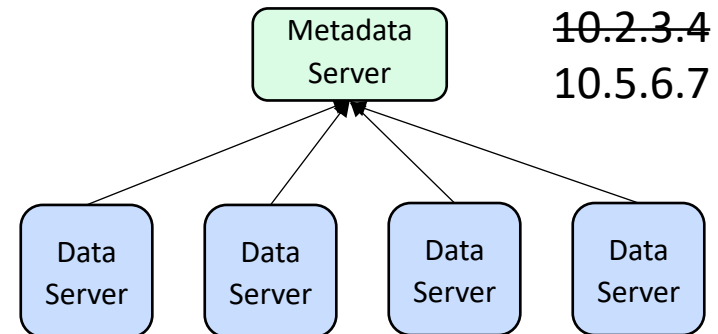
Custom Controller – Leader Solution



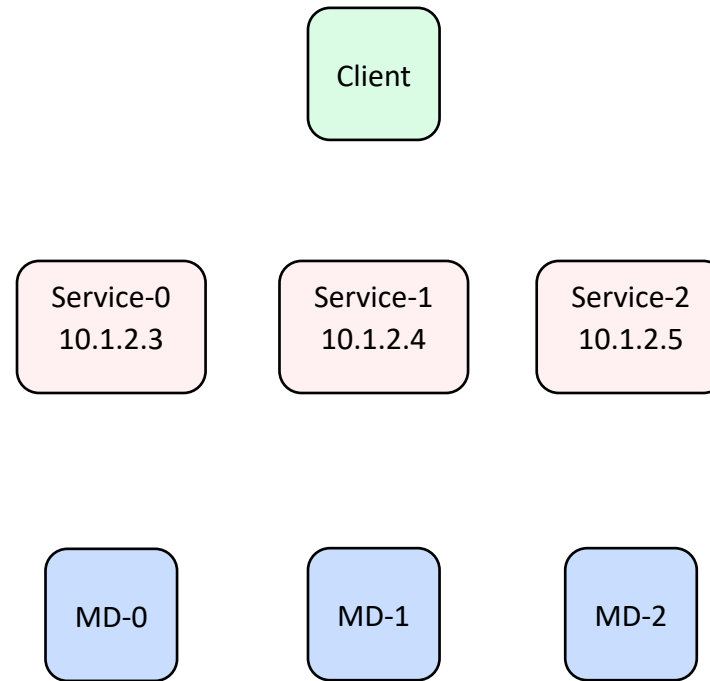
Hadoop/HBase DNS Caching Bugs



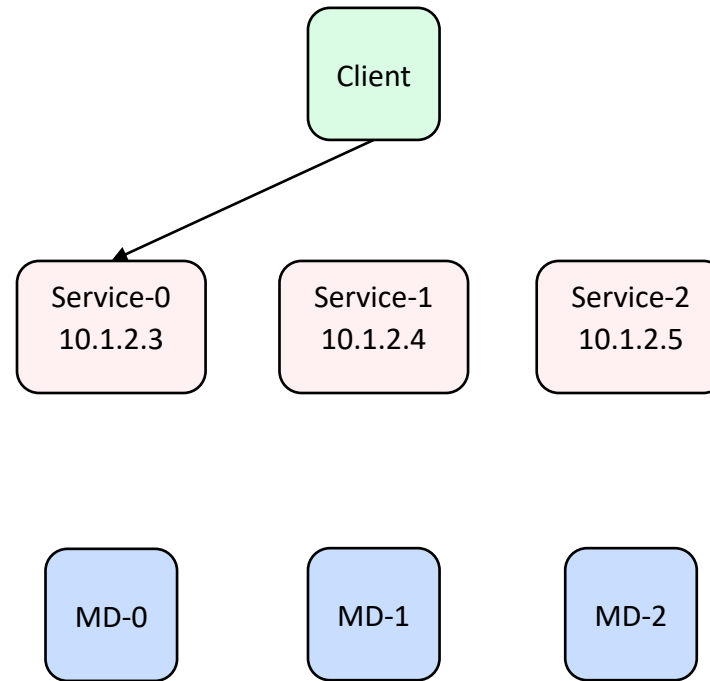
- As Pods are recreated, DNS records keep changing
- Hadoop/HBase code caches DNS resolutions, causing failures
- Impact is most severe for metadata servers



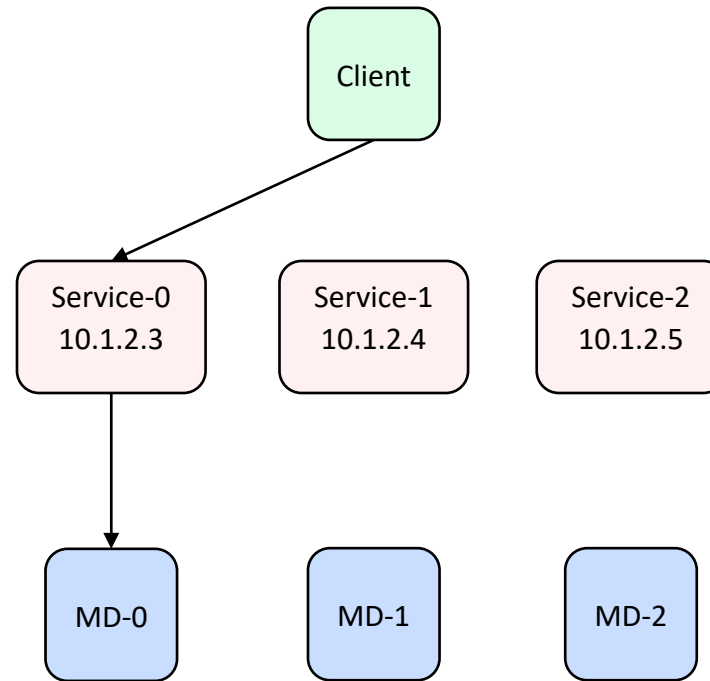
Hadoop Bugs: Interim Solution



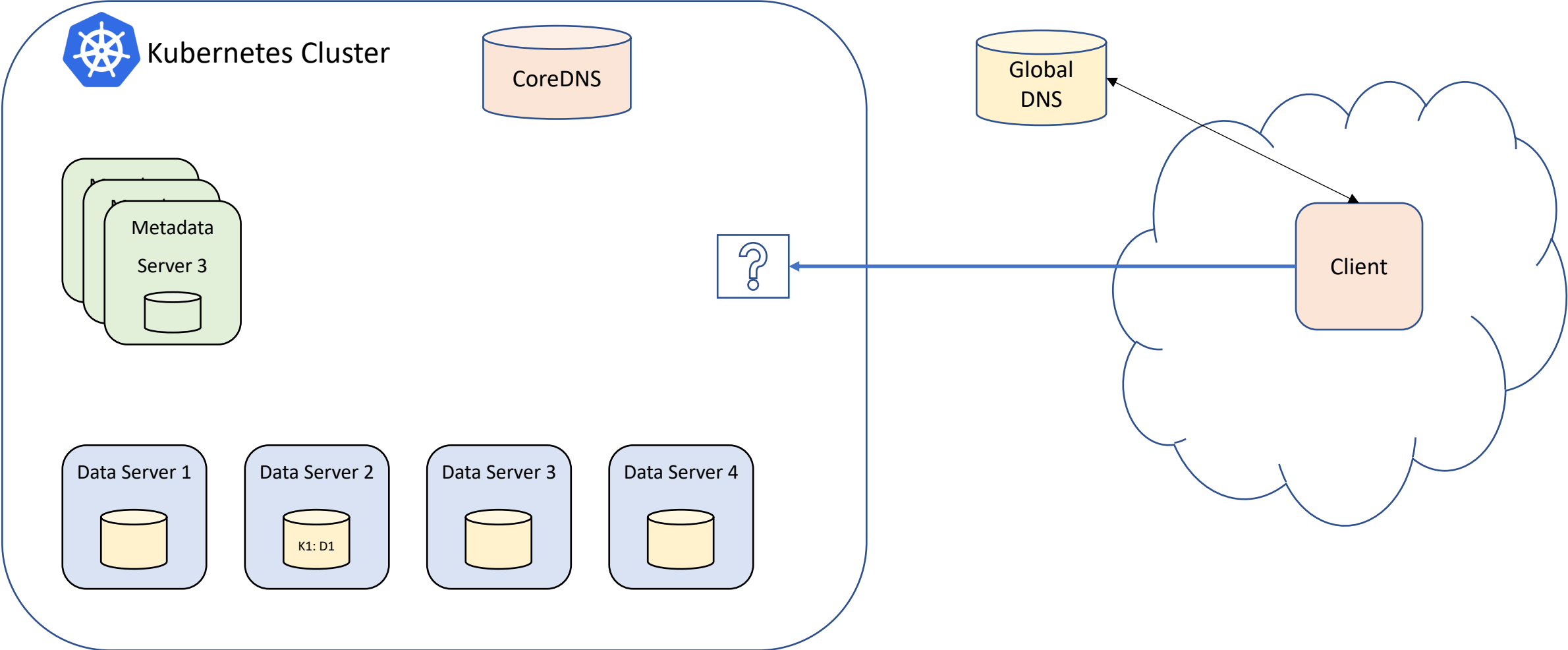
Hadoop Bugs: Interim Solution



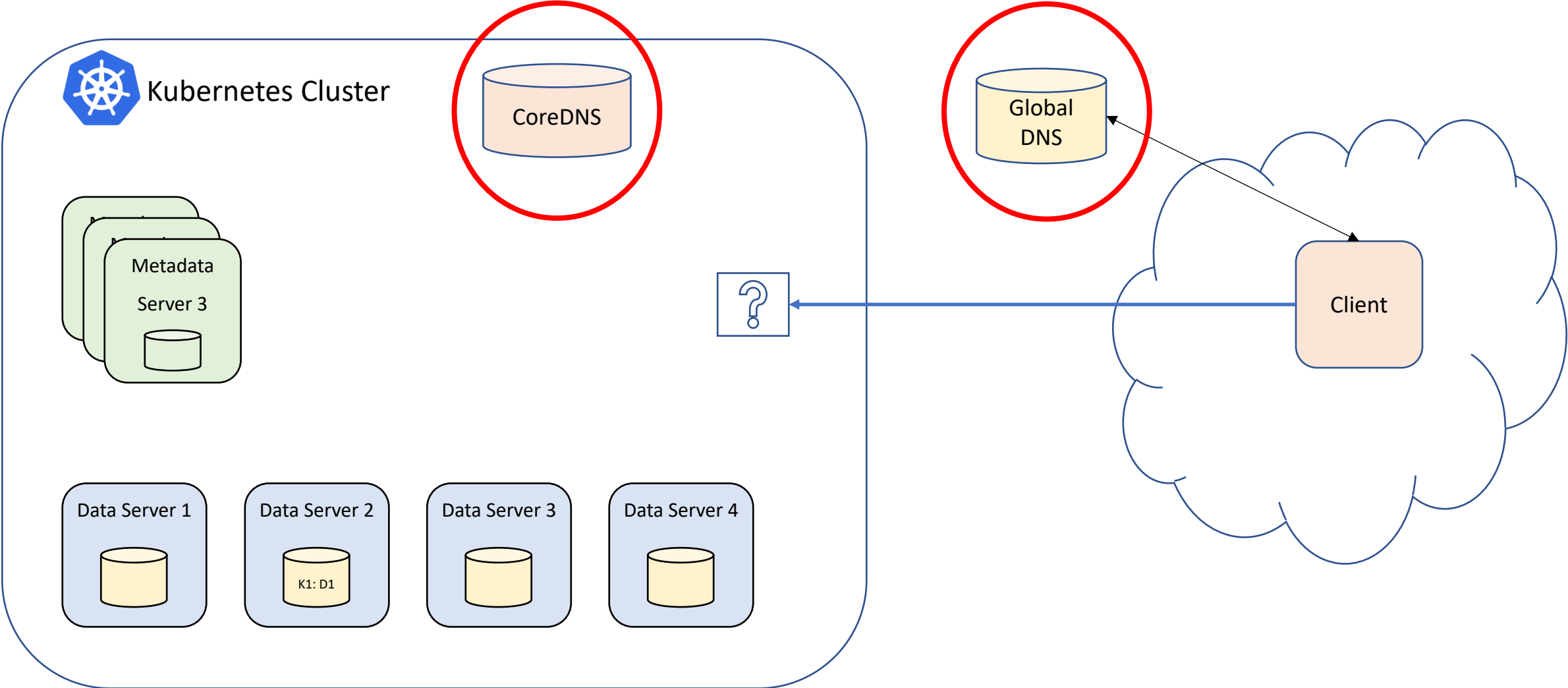
Hadoop Bugs: Interim Solution



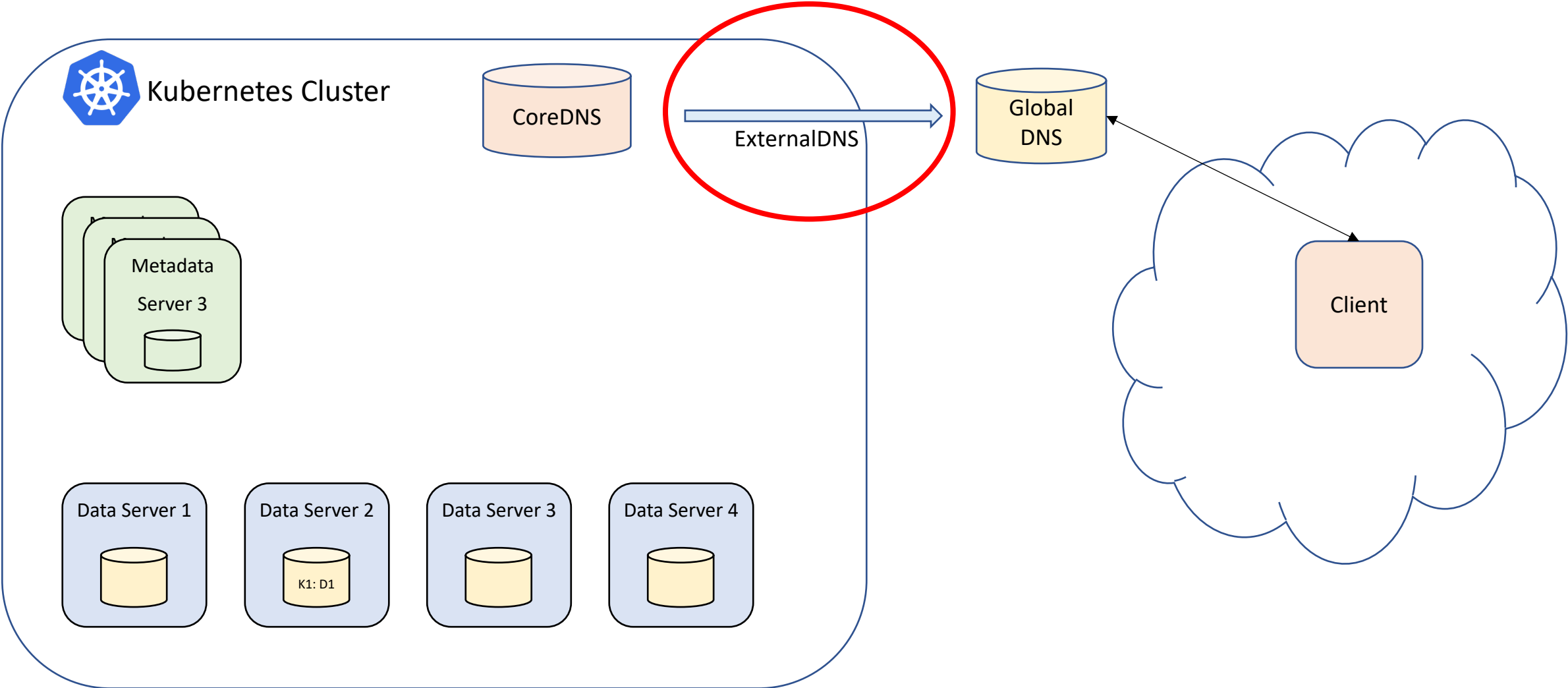
DNS Access



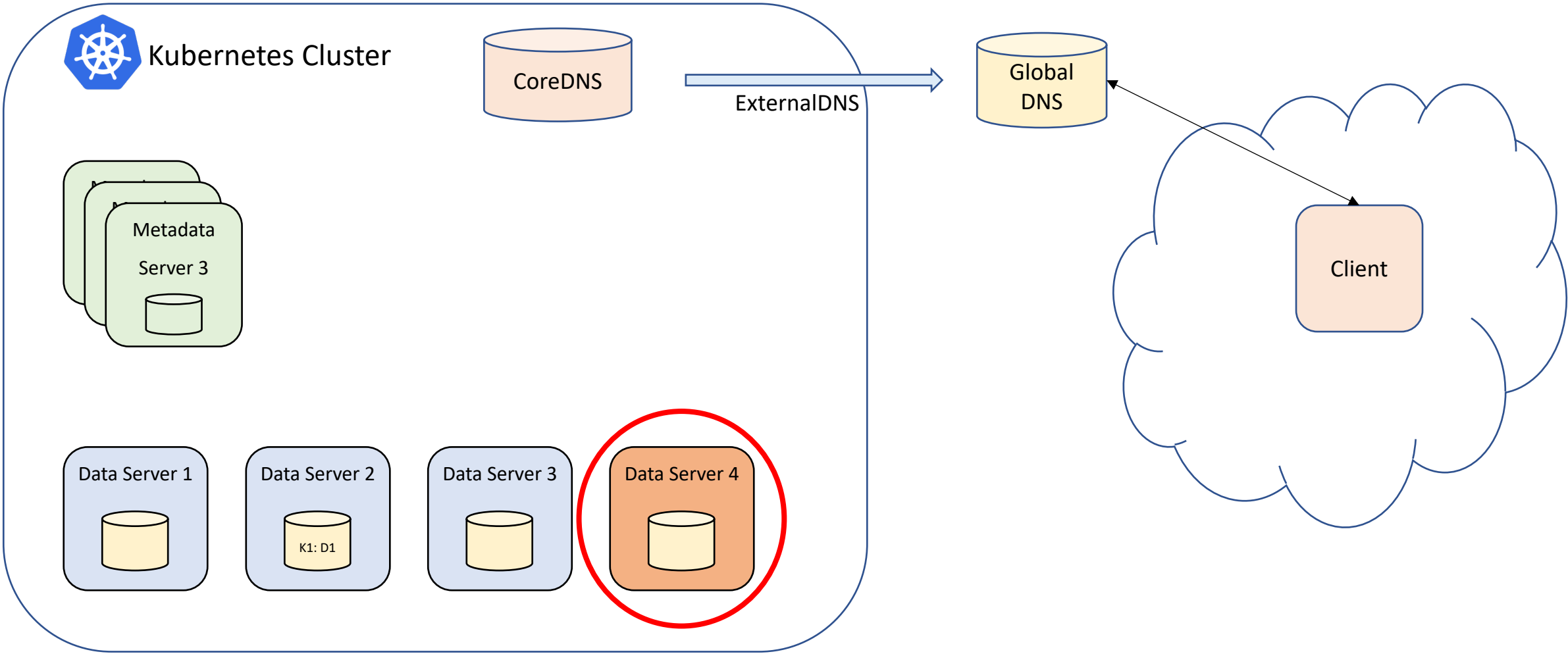
DNS Access



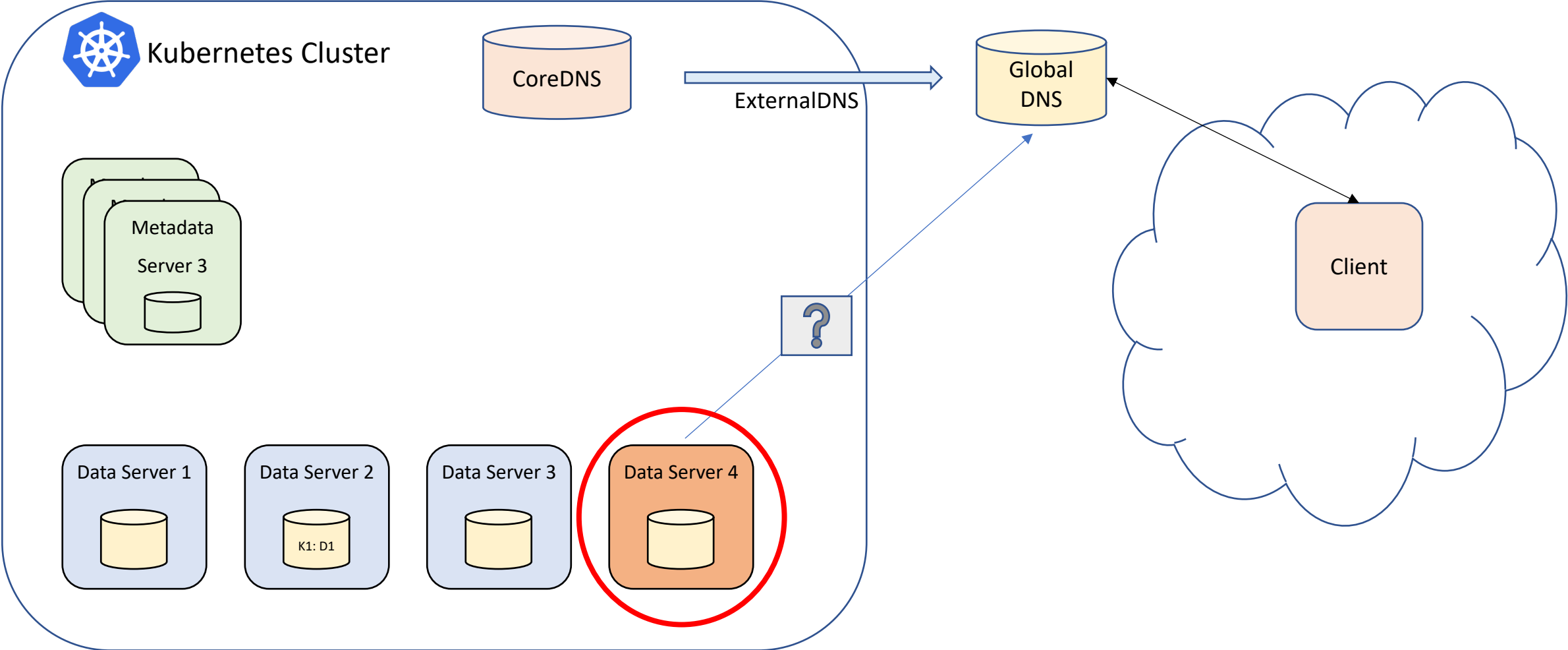
DNS Access



DNS Validation



DNS Validation



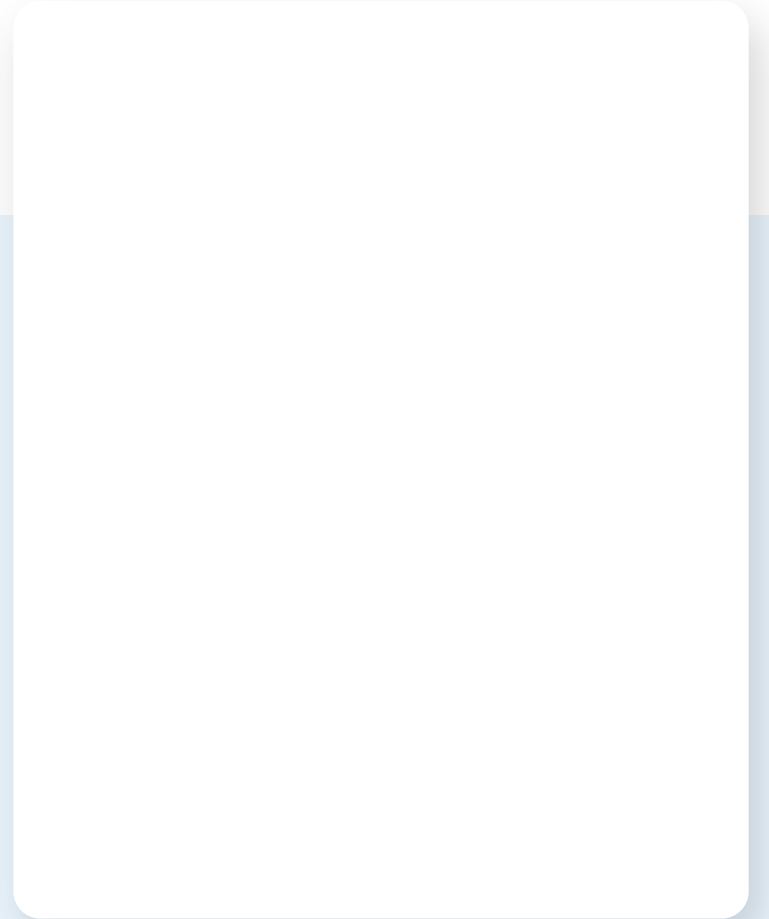
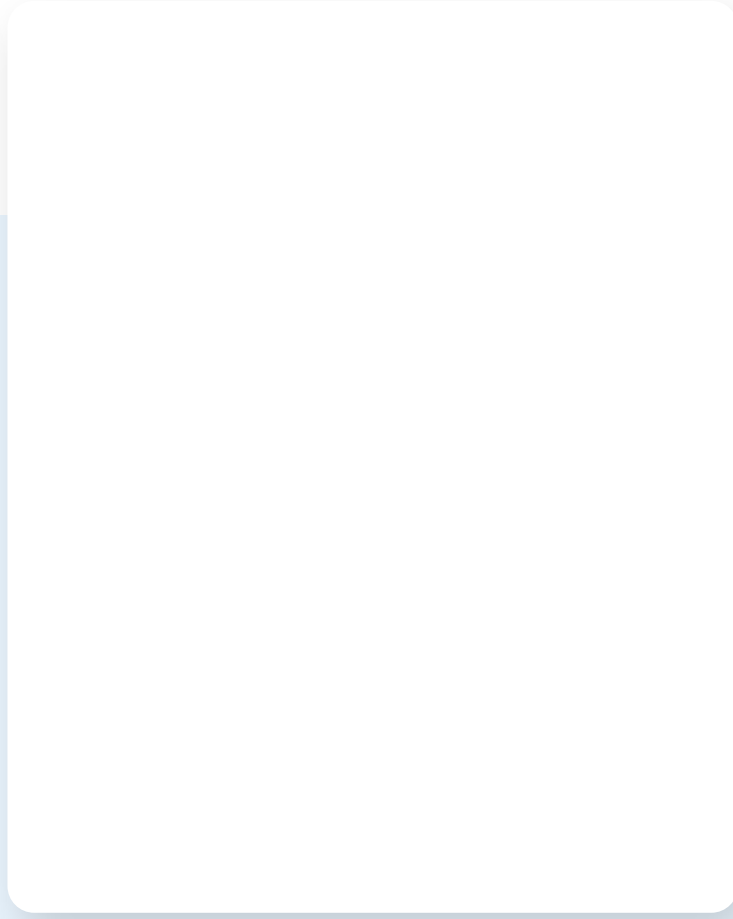
3 AZ Architecture



AZ 1

AZ 2

AZ 3



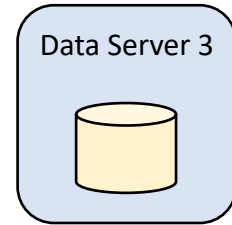
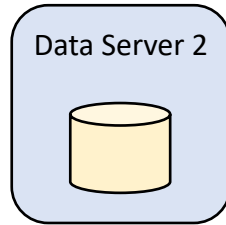
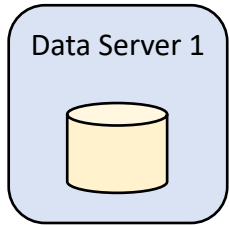
3 AZ Architecture



AZ 1

AZ 2

AZ 3



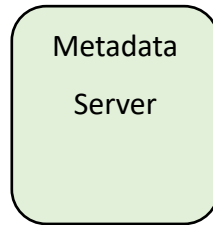
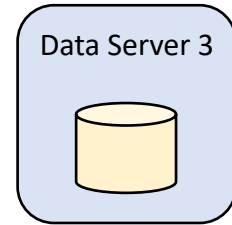
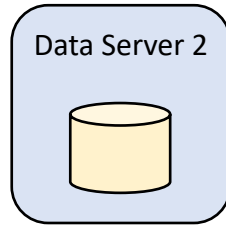
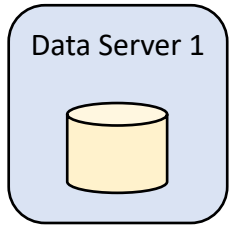
3 AZ Architecture



AZ 1

AZ 2

AZ 3



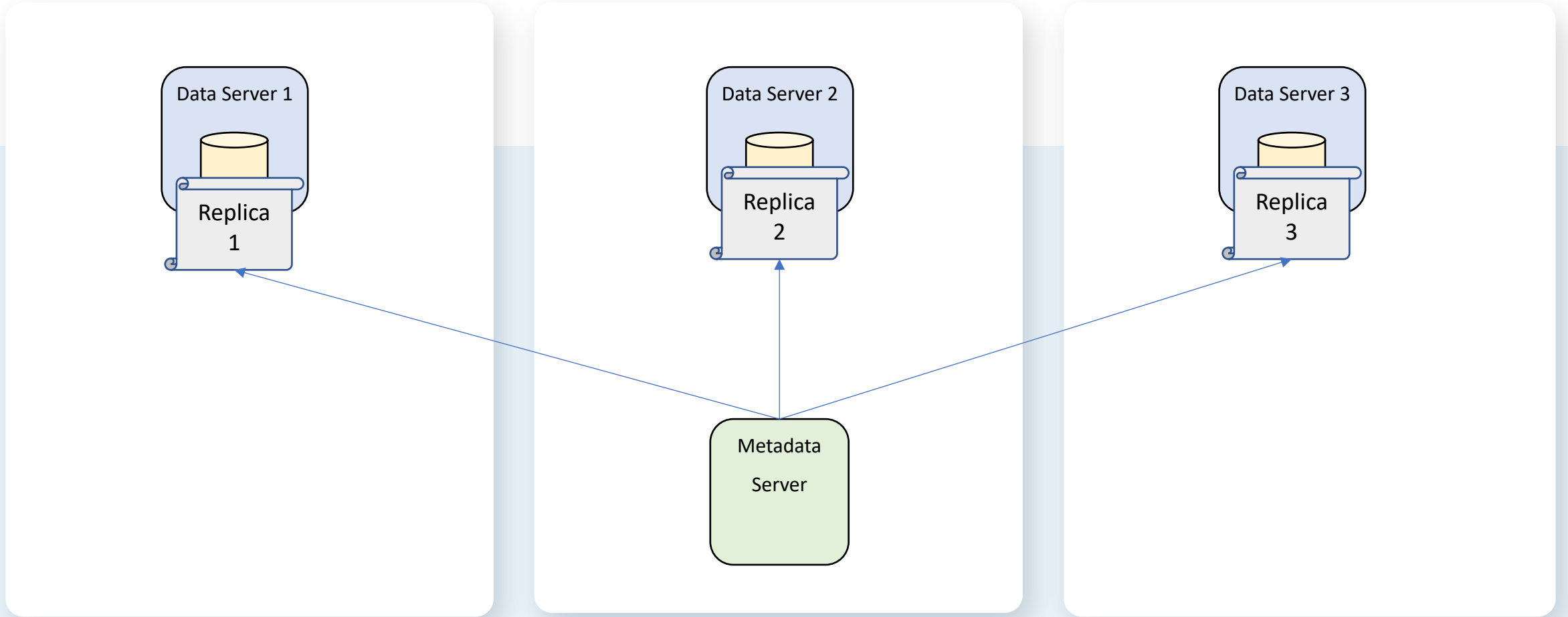
3 AZ Architecture



AZ 1

AZ 2

AZ 3



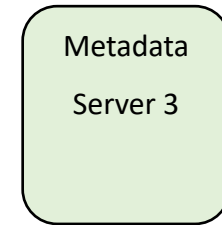
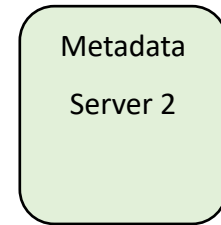
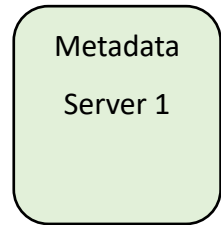
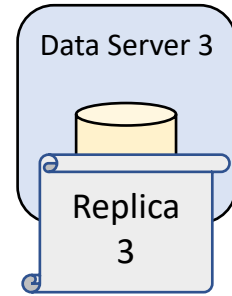
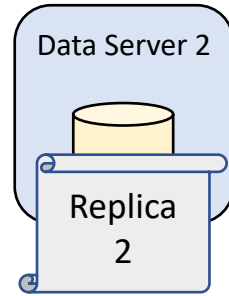
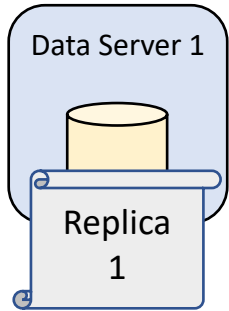
3 AZ Architecture



AZ 1

AZ 2

AZ 3



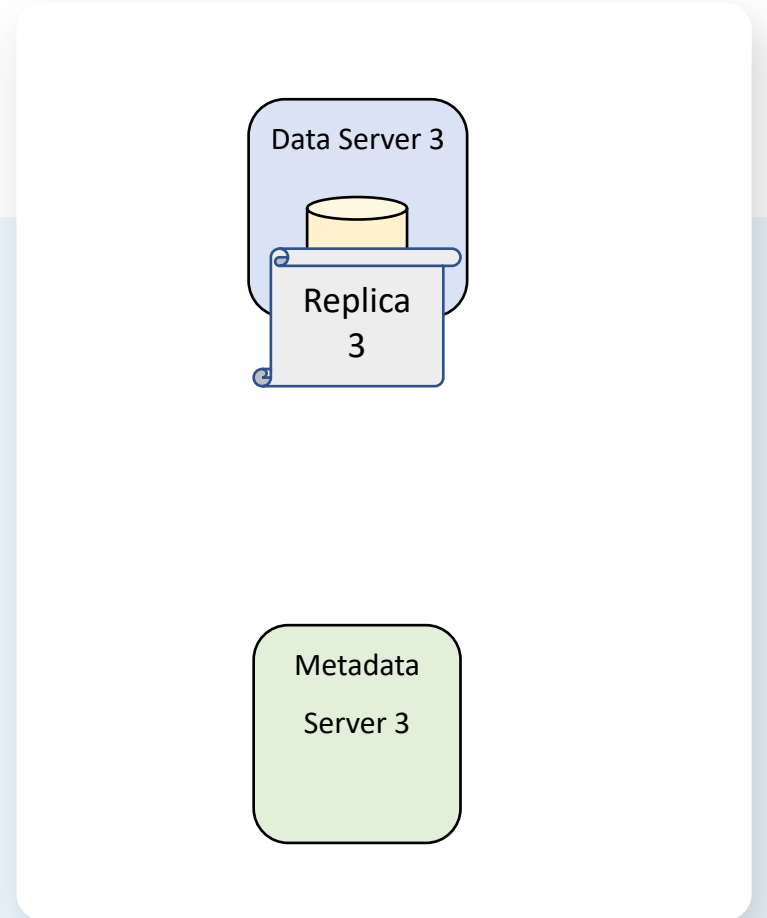
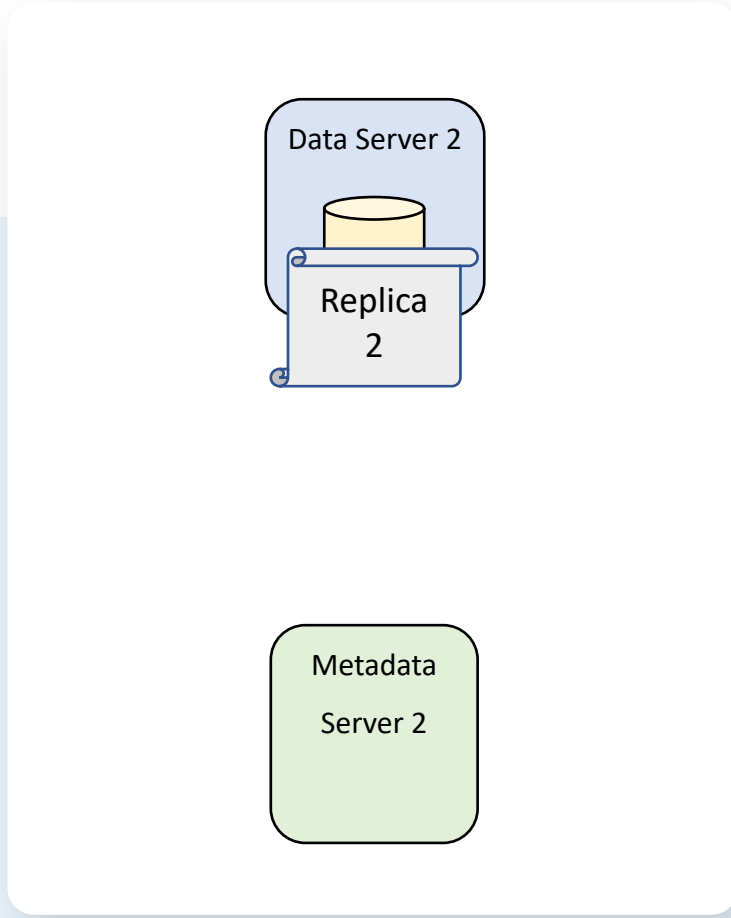
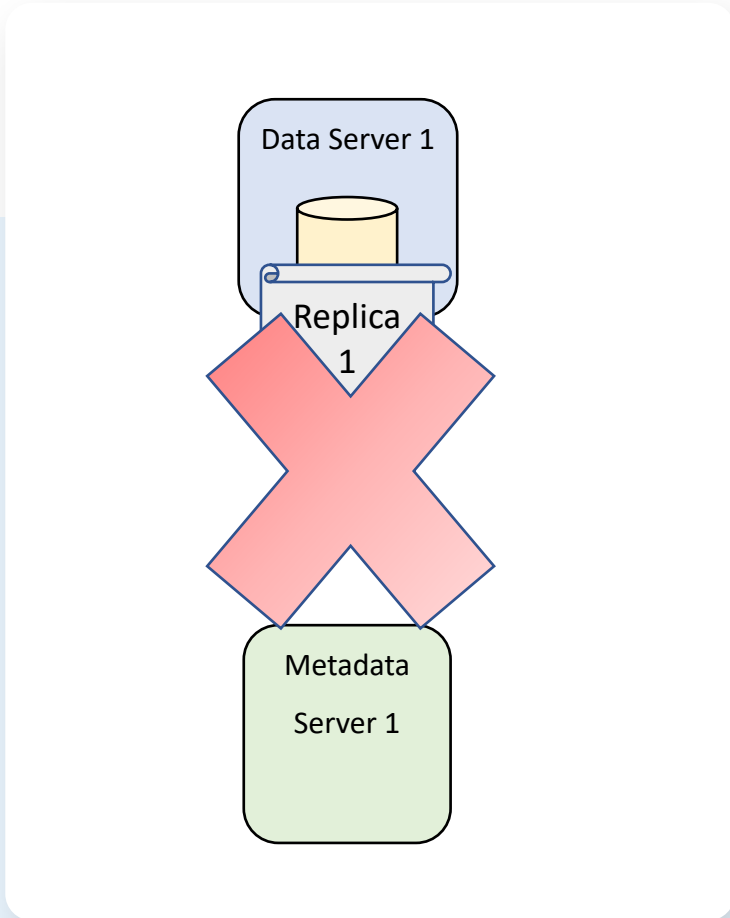
3 AZ Architecture



AZ 1

AZ 2

AZ 3





Thank You